Advances in Clinical and Experimental Medicine

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

2026, Vol. 35, No. 5 (May)

Impact Factor (IF) – 1.9
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. 5
(May)

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. 5 (May)

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

Contents

Reviews

Research letters

Author guidelines in the AI era: Writing for readers, search engines, and reproducibility. Insights from editorial practice

Marek Misiak[1,A–F], Donata Kurpas[2,A,E,F]

[1] Wroclaw Medical University Press, Poland
[2] Division of Research Methodology, Department of Nursing, Faculty of Nursing and Midwifery, 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(5):745–752

Address for correspondence
Marek Misiak
E-mail: marek.misiak@umw.edu.pl

Funding sources
None declared

Conflict of interest
None declared

Received on February 23, 2026
Accepted on March 2, 2026

Published online on April 27, 2026

Abstract

Scientific journals establish author guidelines to ensure manuscript consistency, enhance readability, and maintain editorial standards. However, the rationale behind specific requirements is not always apparent to submitting authors, leading to misunderstandings and noncompliance. This editorial examines the instructions for authors currently applied at *Advances in Clinical and Experimental Medicine* (*Adv Clin Exp Med*), explaining the purpose behind selected regulations that may initially seem arbitrary or overly prescriptive. We analyze requirements concerning manuscript titles (sentence case, study design specification, avoidance of nonstandard abbreviations), author affiliations (institutional hierarchy, geographic formatting), ORCID (Open Researcher and Contributor ID) usage, highlights preparation, taxonomic nomenclature (italicization of genus and species, distinction between genes and proteins), laboratory equipment reporting (manufacturer details, catalog numbers, software versions), abbreviation protocols, and supplementary file management. We demonstrate that these requirements serve essential practical functions: improving search engine optimization and discoverability, ensuring experimental reproducibility, preventing taxonomic and nomenclatural confusion, facilitating rigorous peer review, and enhancing reader comprehension across different formats and access points. The editorial also addresses the evolving nature of author guidelines in the era of artificial intelligence (AI) and digital publishing, emphasizing that editorial policies should function as adaptable documents that respond to technological advances and changing scholarly communication practices. By fostering open dialogue between editors and authors regarding the rationale behind publication requirements, journals can maintain high standards while remaining responsive to the legitimate concerns of the research community. We conclude that transparent communication about editorial policies not only improves compliance but also strengthens the collaborative relationship between journals and the researchers they serve.

Key words: scientific writing, scientific publishing, scientific journal, editorial policies, guidelines as a topic

Cite as
Misiak M, Kurpas D. Author guidelines in the AI era: Writing for readers, search engines, and reproducibility. Insights from editorial practice. *Adv Clin Exp Med*. 2026;35(5):745–752. doi:10.17219/acem/218743

DOI
10.17219/acem/218743

Highlights

- Clear and informative titles improve article visibility, indexing, and reader engagement in scientific publishing.
- Guidelines for authors in *Adv Clin Exp Med* reflect evolving editorial standards shaped by digital discoverability and AI-assisted literature retrieval.
- Understanding the rationale behind manuscript requirements helps authors streamline submission and peer-review processes.
- Continuous adaptation of editorial policies is essential to maintain quality and transparency in modern scientific communication.

Introduction

Most scientific journals formulate specific requirements that must be met by manuscripts submitted to them – as a rule, these are posted on the journal's website, and sometimes links to them can also be found on the journal's social media profiles. If the article is prepared with a specific journal in mind, the submitting author or the whole team should review the journal's instructions for authors before they start writing the manuscript and preparing figures and tables – it can save a lot of work. If the authors choose the place of publication (or at least select the journal) after completing the article, they need to carefully review the preliminary requirements and adapt their manuscript to them.

Some rules provided by editorial offices stem from practical issues – following them facilitates comprehension of articles by both human readers and search engines and AI tools that scan the papers' content. Other regulations are more or less conventional – grounded in editorial traditions or more recent principles widely accepted among editors of scientific journals. Although some requirements stem from editorial traditions, their importance lies in preserving consistency and stylistic coherence, which are central to professional scholarly publishing. Adherence to – and, when necessary, enforcement of – the rules formulated in a given journal is not primarily a question of prestige but of the journal's standing among other journals since many databases and evaluators take into account whether the editors actually perform their duties in the journal in question. In this context, taking such requirements seriously is also in the authors' best interest, since most researchers want to publish in highly esteemed journals.

Editors should explain the reasons behind specific regulations because rules that are understood are more likely to be followed; moreover, it shows that the authors are treated with respect and that the editors do not expect anybody to follow their orders without understanding their rationale but can prove that each tenet in the long run makes both working on the manuscript before publication and reading it following publication easier. It should also be noted that instructions for authors in journals with a similar scope are often significantly different and may even be contradictory when compared – such divergence may be exacerbated by different editorial traditions across various parts of the world (although English serves as the global language of science). Many misunderstandings between authors and editors stem from habit. When experienced researchers submit their manuscripts to journals that follow rules similar to those of other journals, they may not realize that both general and specific requirements are not universally accepted.

Objectives

This editorial outlines selected elements of the Instructions for Authors[1] currently in effect (as of February 2026) at *Advances in Clinical and Experimental Medicine* (*Adv Clin Exp Med*). Where appropriate, explanations are provided for rules whose rationale may not be immediately evident. The 1st aim of the present paper is to help authors interested in submitting manuscripts to our journal understand our requirements and to show them that they are not arbitrary but rooted in widely accepted practices in medical journals and editorial experience. The 2nd aim is to stimulate discussion around the journal's guidelines – they are not written in stone. They should be continually reassessed, as the landscape of scientific publishing is evolving rapidly in the era of AI and editorial rules must evolve accordingly. Examples are provided both from papers published in *Adv Clin Exp Med* and 2 other journals owned by the same publisher – Wroclaw Medical University (Poland): *Dental and Medical Problems* (*Dent Med Probl*) and *Polimery w Medycynie – Polymers in Medicine* (*Polim Med*).

Title of the manuscript

There are 3 rules regarding manuscripts titles in *Adv Clin Exp Med*:

1) Titles should be written in sentence case (only the first word of the title, proper nouns, and genus names are capitalized).

2) The title should include the study design for clinical trials, systematic reviews, or meta-analyses.

3) Other types of reviews should also be defined in the title (e.g., narrative review, scoping review), as determined by Grant and Booth[2] and Ghosh and Choudhury.[3]

If possible, specialist abbreviations in titles should be avoided – for 2 reasons. First, it is always risky to assume that all readers, even specialists in a given field, know the meanings of all abbreviations and acronyms used in publications within their specialty. Second, using too many acronyms (particularly less well-known ones) negatively affects the chances of a paper being highly ranked by search engines (search engine optimization (SEO)). It should be noted that in *Adv Clin Exp Med*, editors often modify the titles of accepted manuscripts to make them more SEO-friendly, and the revised versions are provided to authors for approval during the linguistic editing process.

Only the length of the shorter version of the title (used in pagination) is limited (to 45 characters with spaces); however, it should be borne in mind that overly long titles tend to become convoluted and may overwhelm or confuse readers, while overly short titles may lack essential information. Effective titles balance brevity and completeness, often including the central intervention, outcome, and population studied without unnecessary words.

Formulating an appropriate title for a scientific manuscript may therefore require some reflection, but several studies offer in-depth analysis of this problem, going beyond mere advice. Tullu[4] provided both a classification of papers' titles and suggestions on what a good title should look like in the form of easy-to-comprehend tables. Annesley[5] offered concise advice with many examples. Bahadoran et al.[6] proposed a different classification, with particular emphasis on the length of articles' titles and well-thought-out word choice. Matsubara[7] focused on titles of medical articles – particularly his observations regarding the relationship between titles and manuscript structure are particularly noteworthy. Bavdekar[8] described a proposed procedure for writing a manuscript title and prepared a dedicated checklist based on his insights. Hyland and Zou[9] explored the key features of 5,070 titles in the leading journals of 6 disciplines in the human and physical sciences to identify their typical structural patterns and content; their study does not provide direct recommendations but offers a broader perspective on the issue. Jamali and Nikzad[10] analyzed the relationship between article title type and the number of downloads and citations. Valuable advice has also been offered by Springer Nature,[11] Researchers' Writing Academy,[12] and Multidisciplinary Digital Publishing Institute (MDPI).[13]

In light of the above publications, a good scientific title should be:

– Informative and specific – it should include key elements of the study, such as subject, variables, or outcome;

– Concise and focused – optimal titles are typically short but comprehensive, often recommended at about 10–15 words (though conventions vary by field);

– Precise and accurate – the title must be unambiguous and avoid misleading readers about the content;

– Keyword-rich – include the most critical scientific terms that reflect the study's focus and that researchers would likely use when searching literature;

– Free from jargon and nonstandard abbreviations – these can confuse readers and reduce discoverability through search engines and indexing services;

– Professionally toned – humor, puns, or overly creative phrasing might attract attention in some contexts, but they are usually discouraged in scientific titles because they can reduce clarity and seriousness.

Examples of papers published in our journal with titles formulated according to the rules and suggestions listed above can be divided into 3 groups: 1) articles (mainly clinical trials) with the study design clearly stated in the title – Matys et al.,[14] Yao et al.,[15] and Gao et al.[16]; 2) reviews with a precisely defined type – Jiang and Hou[17] and Szymański et al.[18]; 3) meta-analyses – Qu et al.[19]

Affiliations

A correct affiliation should include 2 components: 1) the name of the institution; 2) the city and country in which the institution is located. If the institution with which an author is affiliated is divided into smaller units (e.g., departments, clinics, institutes, faculties, branches), the name of the specific unit must be provided, rather than only the general name of the institution (university, hospital, etc.). If an author is affiliated with more than 1 entity, each should be listed as a separate affiliation; 2 or more entities should not be merged into a single long affiliation. Names of states or provinces (in the USA, Canada, South Africa, Mexico, Brazil, China, etc.) should not be provided. Names of countries should be given in their shorter forms (e.g., China, not the People's Republic of China or PRC) and in English, even if the name in a given language is also used in English-language sources (e.g., Turkey, not Türkiye). The reason for this rule is to ensure uniformity by providing the names of the same countries in a consistent form, not according to authors' preferences.

Detailed rules regarding affiliations are rare in medical journals – the Medical School of the University of Melbourne (Australia) issued a guideline on affiliations, although it is general in scope (it covers all types of publications, not only journal articles).[20] This issue has been only scarcely studied so far: Khalifa et al.[21] examined different reporting patterns of author affiliations in a cross-sectional evaluation of publications from an Egyptian academic medical institution, while Bachelet et al.[22] offered a protocol for an exploratory case study of author misrepresentation of institutional affiliations.

A fine example of diligently prepared affiliations is a paper by Petrazzuoli et al.,[23] in which the authors were affiliated with 20 different entities. An article

by Martínez-Sabater et al.[24] showcases how to solve an issue frequently encountered involving authors from Spanish- and Portuguese-speaking countries, as well as Italy – multiple scientific and health institutions there have English-language names but are often better known by acronyms of their original (non-English) names – in such a situation, we elect to provide the English name first. The non-English acronym in parentheses is used to ensure the correct identification of a given entity in both languages. In turn, a meta-analysis by Shao et al.[25] exemplifies a problem characteristic of Chinese papers. In this country, a hospital affiliated with a university often forms a single legal entity with that university, and combining them into a single affiliation is not an error. Well-formatted affiliations can also be found in *Dent Med Probl* – e.g., in Yadav et al.[26] and Orzechowska-Wylęgała et al.[27]

ORCIDs

Authors submitting their manuscripts to *Adv Clin Exp Med* are advised to use a unique ORCID number (Open Researcher and Contributor ID). This popular digital tool enables the identification of authors and their research output. A permanent 16-digit identifier distinguishes its holder from other researchers, even if they have a common name, change their surname for any reason, or their name is spelled differently in different publications. It prevents confusion and ensures their work is always correctly attributed. Many journals, publishers, and databases (e.g., Scopus, Web of Science, Crossref) can automatically link their publications to their ORCID profile. An ORCID profile can serve as a portable, standardized academic CV, including publications, grants, affiliations, peer-review activity, and awards.

It should be emphasized here that while the editors make sure, following acceptance of the manuscript, that all authors who have provided ORCIDs are properly indentified, there is no requirement to register in the ORCID database. Using this identifier is a widespread practice in the scientific community, but not an obvious choice, and many renowned researchers choose not to use this identifier; also, in some countries, ORCID registration is less common, particularly among practice-oriented specialists who publish only occasionally and are not full-time researchers. Editors of *Adv Clin Exp Med* respect all motivations for refraining from using this identifier and never inquire about the reasons for not having an ORCID.

Highlights

Authors submitting manuscripts to *Adv Clin Exp Med* are required to begin the main body of the manuscript with the "Highlights" section, which should appear before the Background/Introduction section and provide a concise summary of the most significant findings and the relevance of the article. It aims to help readers quickly understand the key aspects of the study and enhance the article's visibility in scientific databases. Highlights are also helpful in promoting papers on social media and the Internet– in *Adv Clin Exp Med*, they are routinely optimized for search engines using AI tools.

Detailed requirements for highlights are: 1) 3–5 bullet points; 2) maximum 20–25 words per bullet point (avoid complex sentences); 3) short, one-sentence bullet points containing key information; 4) specific, scientific or technical terms should be used instead of abbreviations and jargon (unless they are commonly understood in a given context); 5) active verbs should be used, e.g., 'identify', 'highlight', 'suggest', 'demonstrate'; 6) keywords should be easily understood by search engines (e.g., 'breast cancer relapse', 'PTSD biomarkers').

Examples of appropriately prepared highlights are those in an editorial by Giorgetti et al.,[28] in an original paper by Jędrzejczyk et al.,[29] in a meta-analysis by Fei et al.,[30] and in a review by Lauricella et al.[31] Highlights are also employed in *Polim Med*, like in de Brouwer and Maqsood,[32] and in *Dent Med Probl*, as seen in Pelechá-Salvador et al.[33] or Dąbrowski et al.[34]

Names of species and genes

Names of genus and species should always be provided in italics (e.g., *Homo sapiens*) – this ensures that readers can distinguish between genus/species and:

1) taxonomic names of higher rank (e.g., taxonomic family);
2) viruses (e.g., human papillomavirus); and
3) Latin names of diseases and conditions (e.g., erythema infectiosum).

In the title of the manuscript and at the first mention of an organism in a paper, the name of the genus and species should be written in full. After the first mention, the first letter of the genus name followed by the full species name should be used (e.g., *H. sapiens*), except from situations when the name of the species begins a sentence.

All names of genes should be in italics. The primary reason for this rule is distinguish between genes and proteins.

Strict adherence to rules discussed in this paragraph is crucial in papers dealing with different genera of bacteria – e.g., Erinmez and Zer.[35] The importance of distinguishing between bacteria and disease names is clearly evident in Zhang et al.,[36] while Kim et al.[37] demonstrate that using italics prevents confusion between non-taxonomic plant names and taxonomic nomenclature. In papers published in *Polim Med*, the former is crucial in Adegbolagun et al.[38] and the latter in Sadiq and Ghafil.[39] Discerning between genes and proteins proved paramount, e.g., in Jabłonowska-Babij et al.[40] and Sui and Xi,[41] as well as in Łacina et al.[42] and Rady et al.[43]

Laboratory equipment, reagents, and software

If any equipment, software, reagents, antibodies, or any materials obtained from external sources were used, the following information should be provided:

– exact model of equipment, full name of assay or reagent, catalog number of the antibody, etc.;

– name of the manufacturer;

– location of the manufacturer (country and city (not state or province – e.g., Los Angeles, not California);

– for software – exact version of the software (e.g., IBM SPSS v. 24.0);

– for freeware software (e.g., R) or databases (e.g., Kyoto Encyclopedia of Genes and Genomes (KEGG) or Gene Ontology (GO)), a direct URL for downloading the required version is sufficient.

The rationale for these requirements is clear– only detailed information about used equipment, chemical and biological agents, and software allows for the replication of the described experiments (provided, of course, that the procedures themselves are also clearly outlined). Good examples of such information abundance are papers describing complicated laboratory research – e.g., Rutkowska et al.,[44] Pulathan et al.,[45] Piszko et al.,[46] and Gasztych and Jurczak.[47]

Abbreviations

Rules concerning abbreviations have 2 aims: 1) to enhance clarity – the exact term or name is always denoted by the same abbreviation or acronym; 2) to reduce text length where appropriate. The 1st reason is more important – clarity should never be sacrificed for brevity.

If an abbreviation is introduced in the text, it should be explained when the term or institutional name first appears; then, only the abbreviation should be used, not the full expression (unless the expression occurs at the beginning of the sentence); alternating between the acronym and the full term may confuse some readers. An important rule is that sentences should not begin with an abbreviation or acronym, although if the full name is very long, this rule is not enforced for reasons of brevity and clarity. This is more an editorial convention than a rule with a rational basis. Still, it has been widely accepted in English-language scientific editing for a long time and is also in force in journals published by Wroclaw Medical University.

All abbreviations in the abstract should be explained in the abstract and repeated in the main text, since the abstract and the main text are treated as separate entities (in many instances, only the abstract is read). All abbreviations appearing in a table, figure, or caption must be explained in the legend of a given table/figure, even if they occur also in the main text or in other tables/figures since each table/figure has to be fully comprehensible when viewed outside the context of the whole paper, since some readers focus only on single tables or figures. These practices are intended to accommodate such reading patterns.

In rare cases, when an acronym is more familiar than the expansion, the abbreviation alone can be used (in our journal, this applies to DNA, USA, and SPSS, but not to, e.g., DMSO, DMEM, or PBS). The authors should use standard abbreviations and acronyms, as introducing nonstandard ones can lead to confusion when a new acronym is introduced. Moreover, several combinations of letters are already widely used to denote specific terms (e.g., SD is usually understood as 'standard deviation', OS as 'overall survival', and OD as 'optical density') and should not be employed for other terms, although there are several acronyms with several different possible explanations – that is why only a fraction of abbreviations and acronyms should be seen as truly "obvious", even in a particular context.

All abbreviations used in the text should be explained in the article, not on a separate list – readers should not be forced to refer to a list on another page whenever they encounter a new abbreviation.

Several authors showed how to properly use abbreviations and acronyms in complex papers to improve clarity and avoiding cumbersome, overlong sentences – e.g., Rakotoarison et al.,[48] Li and Ma,[49] or Rogowski et al.[50] This issue is critical in *Polim Med* because papers appearing there frequently describe experiments on various chemical compounds and other substances, as exemplified by Masoom and Khan[51] or Knefel et al.[52] Confusion regarding notation in publications in these research fields could easily lead to a distorted understanding of whole papers.

Supplementary files

Supplementary data are not raw data but any other materials that cannot be published within the paper but are necessary for reviewers and readers to fully understand the study. Therefore, there are 2 reasons for making supplementary data available:

1) Ensuring the integrity of peer review and statistical review since the reviewers have unlimited access to all data required to assess the manuscript fully;

2) Enhancing the informative value of the paper for readers (reproducibility is more related to the shared raw data, which are not discussed here).

There can be both shared and supplementary data for a single manuscript, or only shared or only supplementary data.

The main difference between shared data and supplementary data lies in 3 issues:

1) Shared data are raw data that were analyzed in the paper, while supplementary data are the results of this analysis that, for various reasons, cannot or should not be published within the main body of the text.

2) Shared data are made available for transparency and replicability, while supplementary data are released primarily for clarity and completeness.

3) For original papers and research letters, data sharing is mandatory (with only a few exceptions), while supplementary data are left to authors' discretion; however, in certain situations, peer reviewers or statistical reviewers may request additional data, which may then be included as supplementary files.

There are 3 main reasons for releasing a given set of data as supplementary files instead of including them in the paper itself:

1) The format of the data makes them incompatible with publication in PDF or HTML – e.g., large tables, Excel spreadsheets, highly complex figures that do not fit on a single page and cannot be divided into several smaller items, or datasets in other formats that are not publishable.

2) The number of tables and/or figures exceeds the limit of 10 tables and 10 figures per article, set for manuscripts submitted to *Adv Clin Exp Med*.

3) The authors deem these materials less necessary – including them in the article itself would make it too long; nevertheless, they still contain key information and, in the authors' opinion, should be available to readers.

The supplementary files should be deposited in a repository – public or institutional – openly available for researchers to download; such a repository should offer the option to assign digital object identifiers (DOIs) to the deposited file packages. The editors can deposit the files in the Zenodo repository (https://zenodo.org) if the authors encounter any problems storing the supplementary data themselves. Neither *Adv Clin Exp Med* nor the publisher of this journal (Wroclaw Medical University) is in any way associated with Zenodo – neither financially nor organizationally. The entire dataset should have a single DOI. Each file should be numbered (Supplementary Table 1, Supplementary Fig. 1, etc.), and at least a one-sentence description of the contents of each file should be provided. The DOI and descriptions are then included at the end of the article's main text. Typical errors in this context include a single description for multiple files within a supplementary materials package and a separate DOI for each supplementary file.

Good examples of supplementary files that enhance understanding of a scientific paper are publications by Martínez-Sabater et al.[24] and Doménech-Briz et al.[53]

Supplementary files may seem purely technical; however, Pop and Salzberg[54] showed that, in many cases, supplementary material today is so extensive that it is reviewed superficially or not at all. To address potential issues, several publishers have developed varying levels of guidance regulating their submission for their availability – i.e., Sage Publishing,[55] Springer Nature,[56] Taylor & Francis,[57] Wiley,[58] IOP Publishing,[59] Cambridge University Press,[60] and Dove Medical Press.[61]

Conclusions

Consistent application of instructions for authors supports clarity, reproducibility, and discoverability of scientific work, while also facilitating efficient editorial handling. Moreover, editorial standardization directly contributes to the journal's visibility and positioning in major indexing databases such as Web of Science and Scopus. Consistent formatting, transparent policies, and rigorous enforcement of submission standards enhance credibility, support accurate indexing, and strengthen the journal's reputation and bibliometric performance. In an increasingly metrics-driven environment, clarity and consistency are not merely aesthetic considerations, but strategic elements of sustainable journal development.

Importantly, editorial guidelines should be viewed as "living" documents (i.e., having the capacity to evolve) rather than immutable regulations. Ongoing changes in scientific publishing – particularly those driven by digital platforms, bibliometric evaluation, and AI – require continuous reassessment of editorial standards. Open communication and mutual understanding between authors and editors are crucial to ensuring that journal policies remain relevant, transparent, and supportive of high-quality scientific output. Therefore, authors should should not hesitate to ask questions and share their comments or concerns, particularly concerning rules that, in their view, pose unnecessary challenges and are unjustified obstacles to many researchers. Dialogue between authors and editors may result in:

– providing a clearer explanation of specific requirementsand using more convincing examples;

– reformulating the rules to make them more transparent;

– relaxing specific regulations or even abandoning them entirely (although the latter rarely occurs);

– postponing certain requirements to later stages of manuscript processing (e.g., from initial assessment following submission to after acceptance for publication).

We thus encourage authors to engage in dialogue with us – while we cannot always promise to be flexible, we are always eager to explain and offer support and guidance.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Marek Misiak https://orcid.org/0000-0003-2208-2193
Donata Kurpas https://orcid.org/0000-0002-6996-8920

References

1. *Advances in Clinical and Experimental Medicine*. Instructions for authors. Wrocław, Poland: Wroclaw Medical University Press; 2026. https://advances.umw.edu.pl/en/instructions-for-authors. Accessed February 23, 2026.

2. Grant MJ, Booth A. A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Info Libr J*. 2009;26(2): 91–108. doi:10.1111/j.1471-1842.2009.00848.x
3. Ghosh A, Choudhury S. Understanding different types of review articles: A primer for early career researchers. *Indian J Psychiatry*. 2025;67(5):535–541. doi:10.4103/indianjpsychiatry.indianjpsychiatry_373_25
4. Tullu MS. Writing the title and abstract for a research paper: Being concise, precise, and meticulous is the key. *Saudi J Anaesth*. 2019; 13(Suppl 1):S12–S17. doi:10.4103/sja.SJA_685_18
5. Annesley TM. The title says it all. *Clin Chem*. 2010;56(3):357–360. doi:10.1373/clinchem.2009.141523
6. Bahadoran Z, Mirmiran P, Kashfi K, Ghasemi A. The principles of biomedical scientific writing: Title. *Int J Endocrinol Metab*. 2019;17(4): e98326. doi:10.5812/ijem.98326
7. Matsubara S. Crafting informative titles in medical articles to enhance the comprehension of the study findings. *JMA J*. 2024;7(3):410–414. doi:10.31662/jmaj.2024-0023
8. Bavdekar SB. Formulating the right title for a research article. *J Assoc Physicians India*. 2016;64(2):53–56. PMID:27730781.
9. Hyland K, Zou H. Titles in research articles. *Journal of English for Academic Purposes*. 2022;56:101094. doi:10.1016/j.jeap.2022.101094
10. Jamali HR, Nikzad M. Article title type and its relation with the number of downloads and citations. *Scientometrics*. 2011;88(2):6532773 0781661. doi:10.1007/s11192-011-0412-z
11. Springer Nature. Titles, abstracts & keywords. Berlin, Germany: Nature Publishing Group; 2025. https://www.springernature.com/gp/authors/campaigns/writing-a-manuscript/titles-abstracts-keywords. Accessed February 17, 2026.
12. Clemens A. How to write a title for your research paper: The 10 most common mistakes. Prague, Czech Republic: Researchers' Writing Academy; 2025. https://annaclemens.com/blog/ten-most-common-mistakes-when-choosing-a-paper-title. Accessed February 17, 2026.
13. Bosworth K. How to choose the best title for your research paper. Basel, Switzerland: Multidisciplinary Digital Publishing Institute (MDPI); 2022. https://blog.mdpi.com/2022/03/23/best-title-for-your-research-paper. Accessed February 17, 2026.
14. Matys J, Gedrange T, Dominiak M, Grzech-Leśniak K. Analysis of aerosol generation during Er:YAG laser-assisted caries treatment: A randomized clinical trial. *Adv Clin Exp Med*. 2024;33(10):1087–1095. doi:10.17219/acem/174536
15. Yao B, Chen Y, Chen H, Feng X, Ma X. Study on the role of postoperative rehabilitation based on the ERAS concept for patients undergoing pancreaticoduodenectomy: Protocol for a randomized controlled clinical trial. *Adv Clin Exp Med*. 2024;34(5):821–829. doi:10.17219/acem/189583
16. Gao Y, Wang J, Guo J, Gao J. Psychological distress in Chinese women with benign breast disease and breast cancer during diagnosis: A cross-sectional study. *Adv Clin Exp Med*. 2026;35(2):231–241. doi:10.17219/acem/204102
17. Jiang G, Hou L. Enhancing professional outcome in nursing and midwifery: A systematic review of competence-based education's impact on performance, self-confidence, and anxiety reduction. *Adv Clin Exp Med*. 2026;35(2):333–342. doi:10.17219/acem/202001
18. Szymański M, Skiba MM, Piasecka M. Clinical implications of differences in the pharmacokinetic and pharmacodynamic profiles of ceftriaxone and cefotaxime: A narrative review. *Adv Clin Exp Med*. 2026;35(3):555–562. doi:10.17219/acem/205018
19. Qu Z, Xu C, Qu Z, Liang H. Procedural sedation in emergency departments: Efficacy and adverse outcomes from a systematic review and meta-analysis *Adv Clin Exp Med*. 2026;35(3):419–436. doi:10.17219/acem/205615
20. Melbourne Medical School. Author affiliation guide. Melbourne, Australia: University of Melbourne; 2018. https://medicine.unimelb.edu.au/__data/assets/pdf_file/0003/2126406/Author-Affiliation-Guide.pdf. Accessed December 10, 2025.
21. Khalifa AA, Hussien SM, Ansary EM, El-Gharably AA. Different reporting patterns of author affiliations: A cross-sectional evaluation of publications from an Egyptian medical academic institute. *Turk Med Stud J*. 2023;10(1):13–18. doi:10.4274/tmsj.galenos.2023.2022-5-3
22. Bachelet VC, Uribe FA, Díaz RA, et al. Author misrepresentation of institutional affiliations: Protocol for an exploratory case study. *BMJ Open*. 2019;9(2):e023983. doi:10.1136/bmjopen-2018-023983
23. Petrazzuoli F, Gokdemir O, Antonopoulou M, et al. Resilience of primary healthcare facilities: Experiences from 16 European countries during the COVID-19 pandemic. A mixed-methods study conducted by EURIPA. *Adv Clin Exp Med*. 2025;34(4):487–505. doi:10.17219/acem/194212
24. Martínez-Sabater A, Chover-Sierra E, Pozo-Herce PD, et al. Gender identity stories: Experiences and perspectives of transgender people about healthcare system in Spain. *Adv Clin Exp Med*. 2025;34(12): 2077–2089. doi:10.17219/acem/208133
25. Shao Y, Shao F, Zhou J, Fang S, Zhu J, Li F. The association between hypoglycemia and mortality in sepsis and septic shock: A systematic review and meta-analysis. *Adv Clin Exp Med*. 2024;33(3):197–205. doi:10.17219/acem/166656
26. Yadav M, Rathore N, Katiyar A, et al. Effectiveness of heartfulness meditation in the management of recurrent aphthous ulcers [published online as ahead of print on March 28, 2025]. *Dent Med Probl*. 2025. doi:10.17219/dmp/168912
27. Orzechowska-Wylęgała BE, Wylęgała AA, Zalejska-Fiolka J, Czuba Z, Toborek M. Pro-inflammatory cytokines and antioxidative enzymes as salivary biomarkers of dentofacial infections in children. *Dent Med Probl*. 2024;(6):1027–1034. doi:10.17219/dmp/185733
28. Giorgetti C, Giorgetti A, Boscolo-Berto R. Establishing new boundaries for medical liability: The role of AI as a decision-maker. *Adv Clin Exp Med*. 2025;34(10):1601–1606. doi:10.17219/acem/208596
29. Jędrzejczyk M, Lee CS, Vellone E, et al. Analysis of changes in mental health, cognitive function and self-care behaviors in patients with heart failure: A prospective cohort study. *Adv Clin Exp Med*. 2025; 34(11):1969–1979. doi:10.17219/acem/202773
30. Fei Z, Duan X, Liang J, Sun Z, Tang J. Traumatic complications linked to prophylactic drain placement after hepatectomy: A meta-analysis. *Adv Clin Exp Med*. 2025;34(12):2017–2024. doi:10.17219/acem/201227
31. Lauricella S, Brucchi F, Cirocchi R. The evolution of transanal approaches in rectal cancer surgery. *Adv Clin Exp Med*. 2025;34(12):2187–2196. doi:10.17219/acem/203579
32. De Brouwer H, Maqsood M. Comparison of the optical and mechanical properties of plastics before and after ethylene oxide sterilization. *Polim Med*. 2025;55(2):79–87. doi:10.17219/pim/208083
33. Pelechá-Salvador M, López Domènech S, Márquez-Arrico CF, et al. Endoplasmic reticulum stress in women with polycystic ovary syndrome and gingivitis: A case–control study of metabolic–periodontal interplay [published online as ahead of print on February 9, 2026]. *Dent Med Probl*. 2026. doi:10.17219/dmp/211428
34. Dąbrowski W, Jagiełło K, Zdrojewski TR, Mossakowska M, Chudek J, Górska R. Evaluating changes in dental status among Polish older adults over a decade: A comparative analysis of PolSenior (2009) and PolSenior2 (2019) surveys. *Dent Med Probl*. 2025;(1):23–30. doi:10.17219/dmp/196535
35. Erinmez M, Zer Y. In vitro effects of deferoxamine on antibiotic susceptibility in Gram-negative bacteria. *Adv Clin Exp Med*. 2023;33(5): 491–497. doi:10.17219/acem/169794
36. Zhang T, Zhuo W, Wang WN, Zhao L. Clinical analysis of 338 cases of dacryolithiasis. *Adv Clin Exp Med*. 2025;34(8):1299–1305. doi:10.17219/acem/192223
37. Kim SK, Jayachandran V, Vo TS, Wijesekara I. Safety assessment of turmeric-boswellia-sesame formulation in healthy adult volunteers: An open-label prospective study. *Adv Clin Exp Med*. 2025;34(8): 1277–1288. doi:10.17219/acem/193023
38. Adegbolagun TI, Odeniyi OA, Odeniyi MA. Native and pregelatinized starches of bitter yam as film formers for oral dissolving formulations. *Polim Med*. 2025;55(1):7–19. doi:10.17219/pim/202946
39. Sadiq SI, Ghafil JA. Polyhydroxybutyrate nanoparticle improving the sensitivity of *Pseudomonas aeruginosa* to ceftriaxone and reducing the biofilm formation in vitro. *Polim Med*. 2025;55(1):31–39. doi:10.17219/pim/203765
40. Jabłonowska-Babij P, Majcherek M, Kłopot A, Szeremet A, Wróbel T, Czyż A. The role of *p16* gene and P16INK4a protein in hematologic malignancies and therapeutic implications: A systematic review. *Adv Clin Exp Med*. 2024;34(9):1575–1587. doi:10.17219/acem/192903
41. 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

42. Łacina P, Porzuczek D, Bogunia-Kubik K, Mazur G, Butrym A. Polymorphisms within genes encoding Ikaros family proteins IKZF1 and IKZF3 in multiple myeloma patients treated with thalidomide. *Dent Med Probl*. 2024;61(6):885–892. doi:10.17219/dmp/183776
43. Rady D, Abbass MM, Hakam HM, Rady R, Abou Shady IM. Correlation between the expression of the iNOS, caspase-3 and α-SMA genes in the parotid glands of albino rats following the administration of two antihistamines from two different generations. *Dent Med Probl*. 2024;61(3):335–343. doi:10.17219/dmp/157346
44. Rutkowska E, Kwiecień I, Raniszewska A, et al. Comparison of T cell maturation profiles in the 1st and 5th wave of COVID-19 in the Polish population. *Adv Clin Exp Med*. 2024;34(3):393–405. doi:10.17219/acem/186813
45. Pulathan Z, Ergene ŞM, Altun G, Yuluğ E, Menteşe A. Protective effect of clotrimazole on lung injury in an experimental model of ruptured abdominal aortic aneurysm. *Adv Clin Exp Med*. 2024;34(1):91–100. doi:10.17219/acem/182821
46. Piszko A, Marcula J, Piszko PJ, Nikodem A, Szymonowicz MK, Dobrzyński M. Physico-chemical properties and composition govern adhesion of resin-based dental fissure sealants: A preliminary in vitro study. *Polim Med*. 2025;55(2): 89–103. doi:10.17219/pim/210966
47. Gasztych M, Jurczak N. Rheological properties of pharmaceutical substrates produced with Celugel. *Polim Med*. 2024;54(2):143–147. doi:10.17219/pim/196210
48. Rakotoarison O, Roleder T, Zimoch W, Kuliczkowski W, Reczuch K, Kübler P. Current role of intravascular imaging in percutaneous treatment of calcified coronary lesions. *Adv Clin Exp Med*. 2024;33(11): 1277–1287. doi:10.17219/acem/175273
49. Li D, Ma Y. B10 cells: Development, phenotype, and function in cancer. *Adv Clin Exp Med*. 2024;33(11):1247–1258. doi:10.17219/acem/176378
50. Rogowski Ł, Kowalska J, Bulińska K, et al. Assessment of physical performance and muscle function in hemodialysis patients participating in an exercise regimen: A cluster-randomized, single-center study. *Adv Clin Exp Med*. 2025;34(12):2105–2118. doi:10.17219/acem/202940
51. Masoom M, Khan MA. Efficacy of sulforaphane in skin cancer animal models: A systematic review. *Polim Med*. 2024;54(2):105–111. doi:10.17219/pim/189406
52. Knefel M, Kantor M, Kępka K, Owczarzy A, Kulig K, Maciążek-Jurczyk M. Applications of bio-printing to promote spinal cord regeneration. *Polim Med*. 2024;54(2):127–134. doi:10.17219/pim/196553
53. Doménech Briz V, Juárez-Vela R, Lewandowski Ł, et al. Sex-related differences in the association of obesity described by emergency medical teams on outcomes in out-of-hospital cardiac arrest patients. *Adv Clin Exp Med*. 2024;34(10):1637–1647. doi:10.17219/acem/193367
54. Pop M, Salzberg SL. Use and mis-use of supplementary material in science publications. *BMC Bioinformatics*. 2015;16:237. doi:10.1186/s12859-015-0668-z
55. Sage Publishing. Supplementary files on SAGE journals: Guidelines for authors. Thousand Oaks, USA: Sage Publishing; 2025. https://journals.sagepub.com/pb-assets/cmscontent/TPS/Supplemental_data_on_sjo_guidelines_for_authors%20(2)-1615455093.pdf. Accessed December 10, 2025.
56. Springer Nature. Supplementary information. Berlin, Germany: Nature Publishing Group; 2025. https://www.nature.com/nature/for-authors/supp-info. Accessed December 10, 2025.
57. Taylor & Francis. Enhancing your article with supplementary material. Abingdon-on-Thames, UK: Taylor & Francis; 2025. https://authorservices.taylorandfrancis.com/publishing-your-research/writing-your-paper/enhancing-your-article-with-supplementary-material. Accessed December 10, 2025.
58. Wiley. Supporting information FAQs. Hoboken, USA: Wiley; 2025. https://authorservices.wiley.com/author-resources/Journal-Authors/Prepare/manuscript-preparation-guidelines.html/supporting-information.html. Accessed December 10, 2025.
59. IOP Science. Supplementary material and data in journal articles. Bristol, UK: IOP Publishing; 2025. https://publishingsupport.iopscience.iop.org/questions/supplementary-material-and-data-in-journal-articles. Accessed December 10, 2025.
60. Cambridge University Press. Publishing supplementary material. Cambridge, UK: Cambridge University Press; 2025. https://www.cambridge.org/core/services/authors/publishing-supplementary-material. Accessed December 10, 2025.
61. Dove Medical Press. Supplementary data. Macclesfield, UK: Dove Medical Press (Taylor & Francis); 2025. https://www.dovepress.com/author-guidelines/supplementary-data. Accessed December 10, 2025.

Platelet-rich plasma plus hyaluronic acid vs platelet-rich plasma intra-articular injections for knee osteoarthritis: A systematic review and meta-analysis

Marcin Tomaszewski[1,A–F], Jacek Smereka[2,E,F], Mahdi Al-Jeabory[1,3,C–F], Piotr Fudalej[1,4,B,C,E,F], Michał Pruc[5,B–D,F], Maciej W. Krupowies[1,E,F], Zuzanna Gaca[1,D–F], Łukasz Szarpak[1,5,A–F]

[1] Department of Clinical Research and Development, LUXMED Group, Warsaw, Poland
[2] Department of Emergency Medical Service, Wroclaw Medical University, Poland
[3] Department of Trauma and Orthopedic Surgery, Ruda Śląska City Hospital, Poland
[4] ORTHOS Multi-Specialty Hospital Komorowice, LUXMED Group, Poland
[5] Institute of Medical Science, Collegium Medicum, The John Paul II Catholic 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(5):753–764

Address for correspondence
Łukasz Szarpak
E-mail: lukasz.szarpak@kul.pl

Funding sources
None declared

Conflict of interest
None declared

Received on January 15, 2025
Reviewed on March 18, 2025
Accepted on August 13, 2025

Published online on May 11, 2026

Cite as
Tomaszewski M, Smereka J, Al-Jeabory M, et al. Platelet-rich plasma plus hyaluronic acid vs platelet-rich plasma intra-articular injections: The effect on pain and physical function in the treatment of knee osteoarthritis. Systematic review and meta-analysis. *Adv Clin Exp Med*. 2026;35(5):753–764. doi:10.17219/acem/209523

DOI
10.17219/acem/209523

Abstract

Background. Knee osteoarthritis (KOA) is a common condition characterized by pain, stiffness, and reduced function, significantly impacting quality of life.

Objectives. This study aimed to evaluate the combined use of platelet-rich plasma (PRP) and hyaluronic acid (HA) compared to PRP alone in treating KOA. The hypothesis was that PRP + HA would provide superior outcomes in pain relief and functional improvement due to their complementary biological effects.

Materials and methods. A systematic review and meta-analysis were conducted according to PRISMA guidelines. The analysis included 11 trials with a total of 892 participants. The studies compared the efficacy of PRP + HA compared to PRP alone. Key outcomes evaluated were changes in the visual analogue scale (VAS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), and Inter-national Knee Documentation Committee (IKDC) scores at 1, 3, 6, 12, and 24 months post-treatment.

Results. At baseline, no significant differences in VAS scores were observed between the 2 groups (5.82 ±2.71 for PRP + HA vs 5.66 ±2.96 for PRP). By the 6-month follow-up, PRP + HA showed a more substantial reduction in VAS scores (−2.83 ±1.60 vs −2.56 ±1.66; standardized mean difference (SMD) = −1.08, 95% confidence interval (95% CI): −2.22 to 0.05; $p < 0.001$), with the largest improvement seen at 24 months (−2.7 ±0.2 for PRP + HA vs 0.4 ±0.23 for PRP; SMD = −14.10, 95% CI: −17.41 to −10.79; $p < 0.001$). WOMAC scores at 12 months also favored the PRP + HA group (−25.26 ±15.24 vs −19.6 ±14.20; SMD = −0.95; $p = 0.01$). IKDC scores showed no significant differences between groups.

Conclusions. The combination of PRP and HA provides superior pain relief and functional improvement compared to PRP alone, particularly at 6 and 24 months. These findings support the inclusion of PRP + HA in KOA treatment protocols and warrant further investigation into its long-term benefits.

Key words: meta-analysis, knee, hyaluronic acid (HA), platelet-rich plasma (PRP), knee osteoarthritis (KOA)

Highlights

- The combination of platelet-rich plasma (PRP) and hyaluronic acid (HA) significantly improves pain relief and physical function in knee osteoarthritis (KOA) compared to PRP alone.
- PRP + HA therapy shows superior results at 6 and 24 months, demonstrating its prolonged efficacy in managing KOA symptoms.
- The meta-analysis underscores the synergistic biological effects of PRP and HA, including anti-inflammatory and chondroprotective mechanisms.
- PRP + HA provides a safe and effective alternative to current treatments, with minimal adverse effects and no systemic complications.
- Results support the integration of PRP + HA into clinical protocols for KOA, particularly for long-term management.

Introduction

Osteoarthritis (OA) of the knee joint (KOA) is one of the most common causes of impaired joint function and reduced quality of life in older adults; it is 4 times more common in women than in men.[1–3] As people age and the prevalence of overweight increases, joints are subjected to greater stress. This, along with ongoing biomechanical and biochemical processes, leads to damage of the joint cartilage and subchondral bone in a manner that cannot be reversed. Other risk factors for OA include anatomical factors, obesity, and a previous history of trauma. This condition negatively impacts quality of life, contributing to poor overall health.[1,4,5]

To help relieve the symptoms of OA, patients suffering from joint pain, swelling, stiffness, and a decline in mobility, resulting in partial disability and requiring assistance with daily activities, need medications. These are difficult to formulate, as in addition to managing pain, slowing disease progression and improving knee function are the main goals.[1,5–7] Treatment options include surgical and non-surgical interventions. Non-surgical treatment is recommended for patients with early stages of KOA (Kellgren–Lawrence (KL) grade I–III), while surgery is performed in patients with advanced KOA, in whom conservative treatment has not been effective. Surgery provides definitive treatment for KOA (KL grade IV).[3,8] The most common methods of treating KOA include chronic use of nonsteroidal anti-inflammatory drugs (NSAIDs), intra-articular injections, limb offloading, rehabilitation, and physical therapy.[6,7] These methods relieve symptoms but do not stop disease progression or repair damaged tissues.

The field of orthobiological treatments for osteoarthritis is broad, with several therapies demonstrating potential in decreasing inflammation, supporting tissue regeneration, and enhancing joint functionality.[9] Blaga et al. highlight numerous clinical benefits of platelet-rich plasma (PRP), including its minimally invasive nature, favorable safety profile, and the fact that most adverse events are related to the injection procedure itself rather than the substance. Additionally, PRP promotes faster recovery and does not trigger an immune response.[10] Soufan et al. also pointed out that both intra-articularly administered PRP and hyaluronic acid (HA) did not exhibit systemic side effects, unlike anti-inflammatory drugs.[11] It is crucial to recognize that the efficacy of treatment outcomes is likely affected by variables such as the composition of the product, the dosage administered, and the method of application, with these factors varying according to the specific classification of the disease being addressed.[12] Finally, it is essential to consider that therapies such as PRP may encounter challenges related to regulatory compliance under Good Manufacturing Practice (GMP) standards, which can hinder the integration of these solutions into routine clinical practice.

Among the intra-articular injections, HA and PRP obtained from the patient's blood are widely used. Both HA, which is a natural component of synovial fluid and influences the synthesis of proteoglycans and glycosaminoglycans, and PRP,[13,14] rich in platelets, leukocytes, growth factors, cytokines and other plasma proteins,[15] have anti-inflammatory and chondroprotective effects, contributing to pain reduction and improved joint function.[16,17] Hyaluronic acid improves viscoelasticity, increasing shock absorption, reduces synovial inflammation and protects cartilage. Hyaluronic acid can bind to the CD44 receptor of chondrocytes and reduce interleukin (IL)-1β expression, which decreases activity of matrix metalloproteinases (MMPs): 1, 2, 3, 9 and 13. Platelet-rich plasma, rich in growth factors, modulates the inflammatory response, regulate sangiogenesis and cell differentiation, and stimulates endogenous HA synthesis.[3] Platelet-rich plasma and HA have shown effectiveness in reducing KOA symptoms and improving joint function in many randomized controlled trials (RCTs) and systematic reviews.

Hyaluronic acid predominantly reinstates the viscoelastic characteristics of synovial fluid, alleviating pain via mechanical support for cartilage, whereas PRP promotes chondrocyte proliferation, cartilage formation, and endogenous HA synthesis; postinjection adverse effects are generally minimal, encompassing transient pain, swelling, or minor joint limitations that subside within days.[1]

Despite numerous studies supporting the effectiveness and safety of HA and PRP individually, research on the simultaneous use of both treatment methods remains scarce. Investigating this combination is essential for elucidating the potential synergistic benefits and guiding clinical practice. The combined action of PRP and HA, through distinct yet complementary mechanisms, could have a synergistic effect: PRP promotes tissue repair and regeneration, while HA improves joint lubrication and protects cartilage.[18] This study seeks to provide a robust evidence base for the use of PRP + HA in the treatment of KOA, potentially offering a more effective alternative to current management strategies. By advancing our understanding of the role of biological therapies in treating degenerative joint diseases, this study could significantly impact patient care and quality of life for those suffering from KOA.

Objectives

The aim of this study was to evaluate the efficacy of combined PRP and HA treatment in treating KOA, focusing on pain relief, functional improvement, and inflammation reduction compared to PRP alone. We hypothesized that the synergistic action of PRP and HA, through their distinct biological mechanisms, would result in enhanced therapeutic effects in alleviating KOA symptoms and improving joint function.

Materials und methods

This systematic review and meta-analysis was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines[19] and followed the recommendations outlined in the *Cochrane Handbook for Systematic Reviews of Interventions*.[20] The protocol for this review was registered with the International Prospective Register of Systematic Reviews (Registration ID: CRD42024510649).

Search strategy and study selection

We searched electronic databases including PubMed, CENTRAL (Cochrane Controlled Trials Register), Web of Science, and Embase for relevant studies published between the databases' inception and 24 July 2024. The following search strategy was used to optimize both search specificity and sensitivity, with the following keywords and phrases: "Platelet Rich Plasma" OR "Plasma, Platelet-Rich" OR "PRP" AND "Hyaluronic Acid" OR "HA" OR "Hyaluronate, Sodium" OR "Hyaluronate Sodium" AND "knee" AND "osteoarthritis" OR "osteoarthritides" OR "osteoarthr*" OR "arthrosis" OR "degenerative arthr*" OR "knee OA".

The search was limited to studies published in English. Additional studies were identified by manually searching the reference lists of all eligible studies and relevant review articles. Where necessary, authors were contacted via email to request additional information or clarifications. Two researchers (M.T. and M.A-J.) conducted the literature search separately, and any discrepancies were resolved through consensus, with the senior author (Ł.S.) facilitating discussions.

All search records were independently screened for eligibility by 2 authors (M.T. and M.P.). Each author independently documented their findings, providing justifications and comments for each decision. When there was a difference of opinion, the senior author (Ł.S.) facilitated reaching an agreement by engaging in a discussion.

The results were subsequently transferred to the web platform Rayyan (https://www.rayyan.ai), which serves as a repository to simplify the process of eliminating duplicates and independently screening potential records.[21] The initial screening process involved reviewing titles and abstracts and removing repeated studies. Subsequently, the full texts of potentially relevant articles were carefully assessed to determine whether they met the inclusion criteria.

Inclusion and exclusion criteria

We included all studies that met the following evidence-based PICOS criteria: 1) Patients: adult patients with KOA; 2) Intervention: PRP + HA; 3) Comparator: PRP alone; 4) Outcomes: contained complete original data, including at least 1 of the following: visual analogue scale (VAS) score, International Knee Documentation Committee (IKDC) score,[22] and Knee Injury and Osteoarthritis Outcome Score (KOOS), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) score,[23] or adverse events related to the treatment method; 5) Study design: RCTs and non-RCTs. We considered all published literature in addition to selected gray literature and preprints that provided full-length reports.

We used the following exclusion criteria: 1) unpublished data; 2) review studies, congress abstracts, letters to the editor, and clinical trial abstracts; 3) research lacking patient outcome scores; 4) non-human model studies; and 5) basic science studies.

Outcome and data extraction

Relevant data from the included articles were extracted by a pair of independent authors (M.T. and M.P. or B.M. and M.A-J.) into a structured template to ensure consistency and accuracy. If necessary, we communicated with the corresponding authors to clarify details or request raw data on specific endpoints. Any discrepancies were resolved through discussion with all the authors until a consensus was reached. The following data points were

extracted from the included publications: 1) study characteristics (i.e., authors, article title, country, study design, length of follow-up), 2) demographic data of participants (i.e., number of participants, age, sex, body mass index (BMI), severity of KOA); and 3) outcome scores (i.e., VAS, KOOS, IKDC, WOMAC, and adverse events).

For data synthesis across studies, primary outcome measures were documented, including WOMAC and KOOS scores, as well as the VAS at 6 months post-intervention.

Secondary outcomes included the number of patients reporting adverse events.

Outcomes at 1, 3, 6, 12, 18, and 24 months post-intervention were assessed where available. In addition, changes in scores relative to baseline measurements were calculated for each time point.

Risk of bias and quality of evidence assessment

Two reviewers (M.T. and M.A-J.; blinded for review purposes) assessed the quality of the eligible studies independently, and the senior author (Ł.S.) resolved any discrepancies. The Cochrane Risk of Bias tool (ROB 2) was used to assess the following domains: 1) random sequence generation, 2) allocation concealment, 3) blinding of participants and personnel, 4) blinding of outcome assessment, 5) completeness of outcome data, 6) selective outcome reporting, and 7) other potential sources of bias.[24] For non-RCTs, the ROBINS-I tool was used, and the following domains were assessed: 1) confounding, 2) selection of participants, 3) classification of interventions, 4) deviations from the intended interventions, 5) missing data, 6) measurement of outcomes, and 7) selection of the reported results.[25] Each domain was categorized as unclear, low, or high risk. These assessments, both individual and overall, were based on the criteria outlined in the ROB 2 tool.[24]

Data analysis and statistical methods

For data analysis and statistical evaluation, we used RevMan 5.4 software (The Cochrane Collaboration, Oxford, UK) and Stata v. 18 (StataCorp, College Station, USA). The threshold for statistical significance across all study outcomes was set at a p-value of 0.05 (two-sided). We calculated standardized mean differences (SMDs) with 95% confidence intervals (95% CIs) for continuous outcomes and odds ratios (ORs) or risk ratios (RRs) with 95% CIs for dichotomous outcomes, allowing the inclusion of zero-event studies to estimate the pooled effects. In cases where studies presented continuous outcomes via medians and ranges, means and standard deviations (SDs) were calculated using the approach proposed by Hozo et al.[26] Our analysis was based on a random-effects model to account for variations across studies. The Q-value statistical test and I^2 index were used to evaluate heterogeneity. Statistical heterogeneity was classified into 3 categories: high (I^2 > 80%), moderate (I^2 = 25–79%), and low (I^2 < 25%).[27] To detect publication bias, we applied Egger's test and generated funnel plots, focusing on asymmetry in analyses involving more than 10 studies. Sensitivity analysis was conducted using a leave-one-out approach.

Results

Literature search

The search process yielded a total of 3,414 articles (Fig. 1). After removing 2,155 duplicates, 1,163 articles were further excluded after a preliminary review of titles and abstracts, resulting in 96 studies. Among these articles, 85 were excluded based on full-text screening. Finally, 11 publications were selected for meta-analysis.[1,3,6,18,28–34]

The risk of bias for all included studies was assessed as low, as determined using the RoB2 and ROBINS-I tools. A detailed breakdown of the risk of bias assessment is illustrated in Supplementary Fig. 1–4.

Study and patient characteristics

The final analysis involved 11 studies, published between 2015 and 2023, comprising 892 participants, with 452 in the PRP + HA group and 440 in the PRP group. The trials were conducted in China,[28,30,33,34] Italy,[1,6,29] the USA,[18] India,[31] Brazil,[32] and Taiwan.[3] The mean age of the subjects ranged from 46.2 to 62.71 years. There were 8 RCTs and 3 retrospective studies. The follow-up period ranged from 1 day to 24 months. The characteristics of the included studies appear in Table 1.[1,3,6,18,28–34]

VAS score by timeframe

At baseline, there was no significant difference in VAS scores between the PRP + HA and PRP groups (5.82 ±2.71 vs 5.66 ±2.96, respectively; SMD = 0.04; 95% CI: −0.13 to 0.21; p = 0.67; Table 2).

There was also no significant difference in follow-up periods, except at 6 months (2.23 ±2.18 vs 2.51 ±2.25; SMD = −0.41; 95% CI: −0.64 to −0.17; p < 0.001) and 24 months (1.70 ±0.60 vs 4.60 ±1.00; SMD = −3.45; 95% CI: −4.46 to −2.44; p < 0.001) after treatment (Supplementary Fig. 5).

The pooled median change from baseline in the PRP + HA group was significantly higher than in the PRP group at the 24-month follow-up (–2.7 ±0.2 vs 0.4 ±0.23, respectively; SMD = –14.10; 95% CI: –17.41 to –10.79; p < 0.001; Fig. 2).

Sub-analysis of RCTs showed statistically significant differences between the PRP + HA and PRP groups only at 6-month and 24-month follow-ups (2.57 ±2.38 vs 2.87 ±2.0; SMD = –0.49; 95% CI: –0.77 to –0.21; p < 0.001

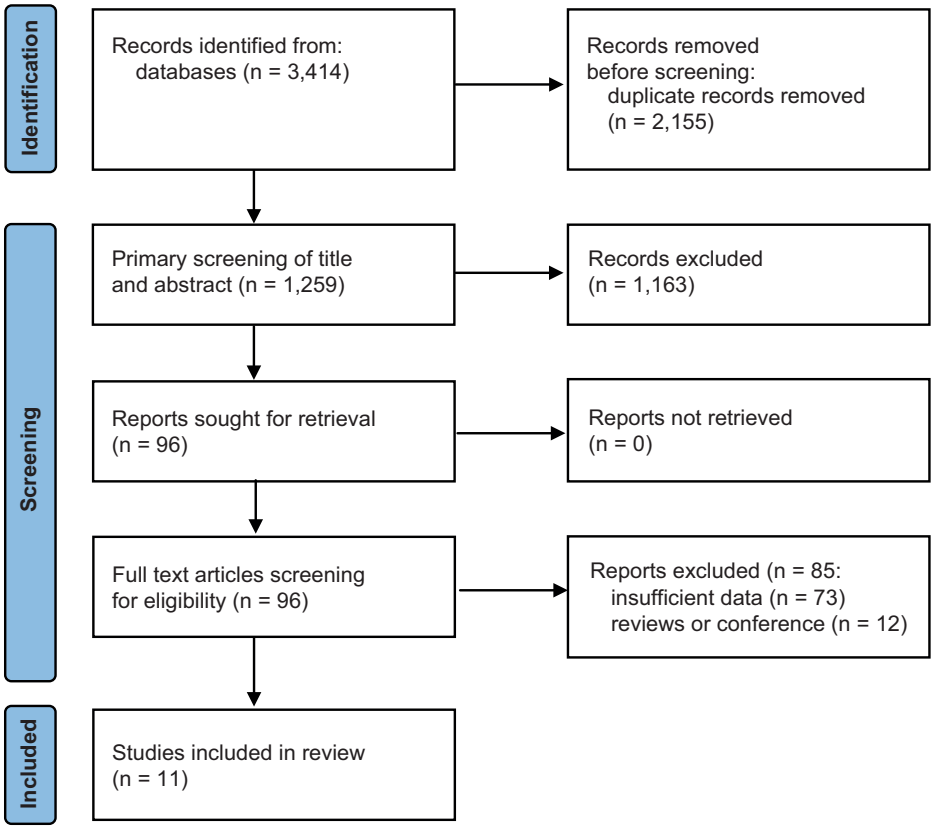

Fig. 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flow diagram

Table 1. Baseline characteristics of included trials

Study	Country	Study design	Study group	Number of participants	Age [years]	Female sex, n (%)	BMI [kg/m^2]	KOA grade (KL)	Follow-up [months]
Abate et al., 2015[1]	Italy	RS	PRP + HA	40	56.7 ±11.2	9 (22.5)	23.7 ±2	II, III	1, 3, 6
			PRP	40	60.9 ±9	19 (47.5)	24.1 ±1.6	II, III	
Wu et al., 2022[3]	Taiwan	RCT	PRP + HA	22	62.2 ±1.5	18 (81.8)	26.3 ±0.9	NS	1, 3, 6, 12
			PRP	23	61.3 ±1.4	18 (78.3)	25.5 ±0.7	NS	
Palco et al., 2021[6]	Italy	RS	PRP + HA	28	62.71 ±7.88	16 (57.1)	NS	II, III	3, 12
			PRP	23	54.04 ±10.4	11 (47.8)	NS	II, III	
Branch et al., 2023[18]	USA	RCT	PRP + HA	32	60.66 ±8.99	20 (62.5)	27.80 ±4.97	I–IV	1, 3, 6, 12, 18, 24
			PRP	32	55.78 ±11.42	16 (50.0)	28.87 ±8.62	I–IV	
Yu et al., 2018[28]	China	RCT	PRP + HA	96	46.5 ±7.5	50 (52.1)	NS	NS	1, 3
			PRP	104	46.2 ±8.6	50 (48.1)	NS	NS	
Ciapini et al., 2023[29]	Italy	RCT	PRP + HA	20	NS	NS	NS	NS	3, 6
			PRP	20	NS	NS	NS	NS	
Guo et al., 2016[30]	China	RS	PRP + HA	63	61.2 ±9.6	45 (71.4)	24.2 ±4.2	I–III	1, 3, 6, 12
			PRP	63	60.7 ±10.1	51 (80.9)	24.6 ±3.9	I–III	
Jacob et al., 2017[31]	India	RCT	PRP + HA	31	NS	NS	NS	NS	6
			PRP	20	NS	NS	NS	NS	
Lana et al., 2016[32]	Brazil	RCT	PRP + HA	33	62 ±6.1	6 (18.2)	29.15 ±7.31	I–III	1, 3, 6, 12
			PRP	36	60.9 ±7.0	67 (19.4)	27.42 ±6.89	I–III	
Xu et al., 2021[33]	China	RCT	PRP + HA	48	57.9 ±4.1	8 (16.7)	21.5 ±2.5	NS	1, 6, 12, 24
			PRP	40	56.9 ±4.2	10 (50.0)	22.5 ±2.3	MS	
Sun et al., 2021[34]	China	RCT	PRP + HA	39	60.6 ±8.4	21 (53.8)	25.0 ±4.6	NS	1, 3, 6
			PRP	39	58.4 ±8.1	17 (43.6)	24.8 ±4.1	NS	

BMI – body mass index; HA – hyaluronic acid; NS – not specified; PRP – platelet-rich plasma; RCT – randomized controlled trial; RS – retrospective study; KOA – knee osteoarthritis; KL – Kellgren–Lawrence grade.

Table 2. Pooled results of different scales among PRP + HA and PRP groups

Outcome	Number of studies	MD (SD)		Events		Heterogeneity between trials		p-value for differences across groups
		PRP + HA	PRP	SMD	95% CI	p-value	I^2 statistics	
VAS score								
Baseline	8	5.82 ±2.71	5.66 ±2.96	0.04	−0.13 to 0.21	0.43	0%	0.67
1 month	3	2.08 ±1.66	2.25 ±2.05	−0.97	−2.73 to 0.79	<0.001	96%	0.28
3 months	4	2.24 ±1.86	2.28 ±1.89	−0.08	−0.33 to 0.17	0.52	0%	0.53
6 months	5	2.23 ±2.18	2.51 ±2.25	−0.41	−0.64 to −0.17	0.52	0%	<0.001
12 months	3	1.64 ±1.24	1.79 ±1.67	−0.19	−0.53 to 0.14	0.25	28%	0.26
24 months	1	1.70 ±0.60	4.60 ±1.00	−3.45	−4.46 to −2.44	N/A	N/A	<0.001
Delta difference baseline (VAS score)								
Δ 1 month	4	−2.54 ±2.07	−2.23 ±1.73	−1.55	−3.32 to 0.21	<0.001	97%	0.08
Δ 3 months	5	−3.00 ±1.83	−1.91 ±1.36	−1.00	−2.25 to 0.24	<0.001	96%	0.11
Δ 6 months	6	−2.83 ±1.60	−2.56 ±1.66	−1.08	−2.22 to 0.05	<0.001	96%	0.06
Δ 12 months	4	−5.1 ±1.4	−4.8 ±1.5	−0.80	−1.98 to 0.37	<0.001	95%	0.18
Δ 24 months	1	−2.7 ±0.2	0.4 ±0.23	−14.1	−17.41 to −10.79	N/A	N/A	<0.001
KOOS score								
Baseline	3	60.07 ±16.44	58.15 ±17.58	0.14	−0.33 to 0.61	0.07	63%	0.57
1 month	2	54.02 ±18.85	52.09 ±15.46	0.13	−0.20 to 0.45	0.96	0%	0.45
3 months	2	65.01 ±18.29	60.28 ±15.45	0.29	−0.04 to 0.62	0.61	0%	0.08
6 months	2	65.0 ±15.39	64.92 ±15.28	0.02	−0.31 to 0.34	0.54	0%	0.92
12 months	2	69.99 ±20.99	70.74 ±19.39	−0.10	−0.47 to 0.27	0.54	0%	0.59
18 months	1	59 ±24.4	59.8 ±19.5	−0.04	−0.53 to 0.45	N/A	N/A	0.89
Delta difference baseline (KOOS score)								
Δ 1 month	2	−7.57 ±15.94	−4.03 ±21.5	−0.24	−0.57 to 0.09	0.44	0%	0.15
Δ 3 months	2	3.42 ±20.31	4.16 ±20.29	−0.05	−0.38 to 0.27	0.68	0%	0.76
Δ 6 months	2	3.41 ±16.64	8.8 ±20.9	−0.29	−0.62 to 0.04	0.78	0%	0.08
Δ 12 months	2	10.33 ±28.18	10.52 ±25.64	0.01	−0.37 to 0.40	0.44	0%	0.95
Δ 18 months	1	−3.7 ±31.8	0.4 ±22.4	−0.15	−0.64 to 0.34	N/A	N/A	0.56
WOMAC (total)								
Baseline	7	46.11 ±26.23	45.15 ±25.18	0.14	−0.17 to 0.46	0.004	69%	0.36
1 month	4	41.3 ±32.03	41.21 ±32.8	−0.05	−0.37 to 0.28	0.21	34%	0.76
3 months	4	23.34 ±15.03	25.29 ±19.94	−0.42	−1.37 to 0.52	<0.001	91%	0.38
6 months	5	33.92 ±28.93	34.57 ±28.91	−0.24	−0.83 to 0.35	<0.001	82%	0.43
12 months	5	24.38 ±24.76	28.33 ±24.95	−0.53	−1.00 to −0.05	<0.001	82%	0.03
18 months	1	27.1 ±23.8	21.0 ±15.1	0.30	−0.19 to 0.80	N/A	N/A	0.23
24 months	2	50.58 ±32.92	54.58 ±38.31	−1.52	−4.77 to 1.74	<0.001	97%	0.36
Delta difference baseline – WOMAC (total)								
Δ 1 month	4	−16.62 ±13.36	−15.01 ±10.1	−0.52	−1.63 to 0.58	<0.001	93%	0.35
Δ 3 months	4	−13.22 ±25.45	−20.64 ±8.66	−0.45	−2.23 to 1.33	<0.001	97%	0.62
Δ 6 months	5	−20.83 ±18.52	−19.42 ±12.71	−0.17	−1.25 to 0.92	<0.001	94%	0.77
Δ 12 months	5	−25.26 ±15.24	−19.6 ±14.20	−0.95	−1.66 to −0.23	<0.001	92%	0.010
Δ 18 months	1	−13.4 ±5.35	−19.0 ±4.24	1.15	0.61 to 1.68	N/A	N/A	<0.001
Δ 24 months	2	−12.76 ±4.59	−8.45 ±7.17	−5.11	−16.04 to 5.83	<0.001	99%	0.36
WOMAC (Pain)								
Baseline	3	9.75 ±6.1	9.49 ±5.04	0.20	−0.15 to 0.55	0.13	51%	0.26
1 month	2	6.1 ±3.84	6.6 ±5.5	−1.04	−4.02 to 1.93	<0.001	98%	0.49
3 months	2	4.59 ±2.9	5.6 ±4.9	−1.59	−5.41 to 2.22	<0.001	98%	0.41
6 months	2	4.79 ±2.86	5.88 ±4.49	−1.40	−4.47 to 1.66	<0.001	97%	0.37
12 months	2	4.16 ±2.86	5.35 ±3.82	−1.34	−3.45 to 0.77	<0.001	96%	0.21

Table 2. Pooled results of different scales among PRP+HA and PRP groups – cont.

Outcome	Number of studies	MD (SD)		Events		Heterogeneity between trials		p-value for differences across groups
		PRP + HA	PRP	SMD	95% CI	p-value	I^2 statistics	
Delta difference baseline – WOMAC (pain)								
Δ 1 month	3	−3.84 ±3.63	−3.45 ±1.95	−0.46	−2.36 to 1.44	<0.001	97%	0.64
Δ 3 months	3	−4.87 ±4.42	−3.98 ±2.46	−0.95	−2.73 to 0.82	<0.001	96%	0.29
Δ 6 months	3	−4.28 ±4.96	−3.81 ±2.98	−0.77	−2.66 to 1.12	<0.001	97%	0.42
Δ 12 months	3	−5.75 ±3.88	−4.55 ±3.15	−1.19	−2.56 to 0.19	<0.001	96%	0.09
WOMAC (stiffness)								
Baseline	4	3.20 ±2.65	2.91 ±2.2	0.50	−0.26 to 1.26	<0.001	91%	0.19
1 month	2	2.86 ±1.95	2.51 ±2.07	0.40	0.05 to 0.76	0.74	0%	0.03
3 months	2	2.06 ±1.39	2.09 ±1.82	−0.38	−1.85 to 1.10	<0.001	93%	0.62
6 months	2	2.07 ±1.12	2.45 ±1.81	−1.32	−4.46 to 1.82	<0.001	98%	0.41
12 months	2	1.16 ±0.91	1.85 ±1.18	−1.52	−3.08 to 0.04	<0.001	93%	0.06
Delta difference baseline – WOMAC (stiffness)								
Δ 1 month	3	−1.37 ±1.2	−1.3 ±0.92	0.40	−2.16 to 2.96	<0.001	98%	0.76
Δ 3 months	3	−1.96 ±1.79	−1.72 ±1.09	0.08	−2.76 to 2.93	<0.001	98%	0.95
Δ 6 months	3	−1.82 ±2.19	−1.4 ±1.03	−0.61	−3.45 to 2.22	<0.001	98%	0.67
Δ 12 months	3	−2.29 ±1.85	−1.48 ±1.43	−1.45	−2.89 to −0.01	<0.001	96%	0.05
WOMAC (function)								
Baseline	4	41.82 ±30.88	39.83 ±30.75	0.37	−0.09 to 0.83	0.005	77%	0.11
1 month	2	23.57 ±15.49	21.87 ±15.55	0.26	−0.09 to 0.62	0.92	0%	0.15
3 months	2	17.32 ±12.05	18.81 ±15.4	−0.72	−2.67 to 1.23	<0.001	96%	0.47
6 months	2	17.41 ±10.95	18.15 ±11.56	−0.36	−1.27 to 0.54	0.02	83%	0.43
12 months	2	13.06 ±9.89	17.64 ±12.1	−0.95	−2.04 to 0.14	0.003	89%	0.09
Delta difference baseline – WOMAC (function)								
Δ 1 month	3	−32.04 ±40.33	−22.24 ±25.25	−0.46	−1.43 to 0.51	<0.001	90%	0.35
Δ 3 months	3	−47.69 ±41.42	−41.44 ±38.25	−0.98	−2.65 to 0.68	<0.001	96%	0.25
Δ 6 months	3	−40.59 ±30.8	−36.35 ±27.82	−1.29	−2.37 to −0.21	<0.001	91%	0.02
Δ 12 months	3	−34.15 ±27.77	−24.87 ±24.59	−2.01	−3.71 to −0.30	<0.001	96%	0.02
IKDC score								
Baseline	2	37.42 ±15.03	37.47 ±10.42	0.02	−0.35 to 0.39	0.69	0%	0.92
1 month	1	53.7 ±14.1	50.1 ±17.3	0.23	−0.27 to 0.72	N/A	N/A	0.37
3 months	1	56.9 ±17.6	50.4 ±17.6	0.36	−0.13 to 0.86	N/A	N/A	0.15
6 months	2	49.41 ±18.04	49.37 ±14.63	0.07	−0.30 to 0.44	0.67	0%	0.71
12 months	1	55.7 ±21.5	50.6 ±19.1	0.25	−0.24 to 0.74	N/A	N/A	0.32
18 months	1	53.4 ±24.5	57.0 ±23.6	−0.15	−0.64 to 0.34	N/A	N/A	0.55
24 months	1	51.8 ±21.9	50.9 ±20.1	0.04	−0.45 to 0.53	N/A	N/A	0.87
Delta difference baseline (IKDC score)								
Δ 1 month	1	15.1 ±19.1	10.9 ±20.5	0.21	−0.45 to 0.53	N/A	N/A	0.87
Δ 3 months	1	18.3 ±22.8	11.2 ±20.7	0.32	−0.17 to 0.82	N/A	N/A	0.20
Δ 6 months	2	12.04 ±19.73	11.9 ±15.4	0.05	−0.32 to 0.42	0%	0%	0.79
Δ 12 months	1	17.1 ±27.4	11.4 ±22.2	0.23	−0.27 to 0.72	N/A	N/A	0.37
Δ 18 months	1	14.8 ±30.3	17.8 ±26.6	−0.10	−0.59 to 0.39	N/A	N/A	0.68
Δ 24 months	1	13.2 ±28.1	11.7 ±22.3	0.06	−0.43 to 0.55	N/A	N/A	0.82

95% CI – 95% confidence interval; N/A – not applicable; HA – hyaluronic acid; NS – not specified; PRP – platelet-rich plasma; RCT – randomized controlled trial; RS – retrospective study; KOA – knee osteoarthritis; SMD – standardized mean difference; MD – mean difference; SD – standard deviation; KOOS – Knee Injury and Osteoarthritis Outcome Score; IKDC – International Knee Documentation Committee; WOMAC – Western Ontario and McMaster Universities Osteoarthritis Index; VAS – visual analogue scale. Results derived from single studies do not constitute a pooled analysis and should be interpreted with caution due to the limited generalizability of such findings.

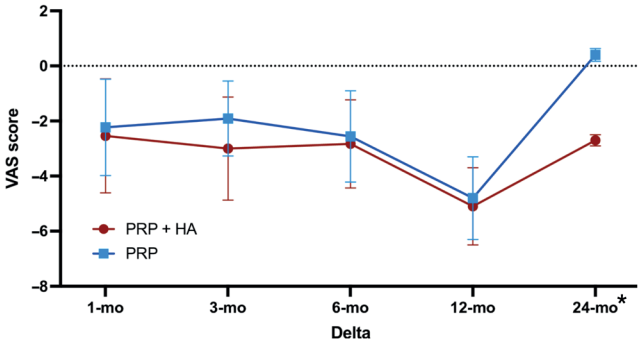

Fig. 2. The evolution of the mean visual analogue scale (VAS) scores changes from baseline in the research groups

Results marked with an asterisk refer to a single study and do not represent a pooled analysis. They should be interpreted with caution, given the limited generalizability of findings based on individual studies.

and 1.70 ±0.60 vs 4.60 ±1.0; SMD = −3.45; 95% CI: −4.46 to −2.44; p < 0.001; Table 2).

Analysis of changes in values using the VAS showed statistically significant differences from baseline only after 24 months (−2.7 ±0.2 for PRP + HA vs 0.4 ±0.23 for PRP; SMD = −14.10; 95% CI: −17.41 to −10.79; p < 0.001; Supplementary Fig. 6).

WOMAC score by timeframe

Baseline WOMAC scores were similar across both PRP + HA and PRP groups (46.11 ±26.23 vs 45.15 ±25.18; SMD = 0.14; 95% CI: −0.17 to 0.46; p = 0.36; Supplementary Fig. 7). All studies were RCTs. Statistically significant differences were noted at the 12-month follow-up, with the PRP + HA group showing a more substantial reduction in WOMAC scores (24.38 ±23.8 vs 28.33 ±24.95; SMD = −0.53; 95% CI: −1.00 to −0.05; p = 0.03). Mean changes between the 12-month follow-up and baseline revealed statistically significant differences in WOMAC score between PRP + HA and PRP groups (−25.26 ±15.24 vs −19.6 ±14.20; SMD = −0.95; 95% CI: −1.66 to −0.23; p = 0.010), as well as between the 18-month follow-up and baseline (−13.4 ±5.35 vs −19.6 ±14.2; SMD = 1.15; 95% CI: −16.04 to 5.83; p < 0.001; Fig. 3).

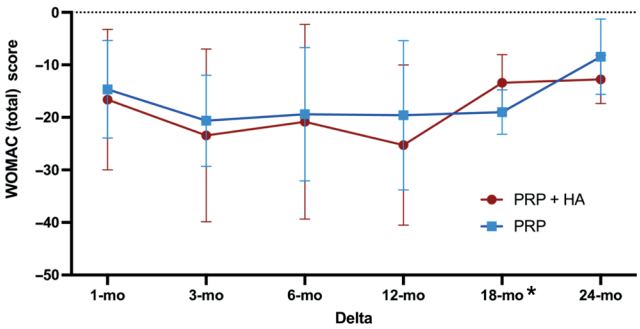

Fig. 3. The evolution of the mean Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scores changes from baseline in the research groups

Detailed WOMAC subscale (pain, stiffness, and function) values are shown in Table 2. A graphic representation of the differences between the measurements and baseline in each subscale is illustrated in Supplementary Fig. 8, Supplementary Fig. 9, and Supplementary Fig. 10, respectively.

KOOS scale by timeframe

Only 3 studies reported KOOS scores in PRP + HA and PRP patient groups. There were no statistically significant differences in the KOOS score either at baseline (60.07 ±16.44 vs 58.15 ±17.58; SMD = 0.14; 95% CI: −0.33 to 0.61; p = 0.57) or during individual follow-ups (1, 3, 6, 12, and 18 months after treatment; p > 0.05; Supplementary Fig. 11). Even when analyzing the differences between individual time points and baseline, no statistically significant differences were noted between the study groups (Fig. 4). Only 1 RCT reported the KOOS scale in the study groups. No statistically significant differences were noted between PRP + HA and PRP at any of the time points (p > 0.05). A similar situation occurred when analyzing the differences in values between individual time points and baseline (Supplementary Fig. 12).

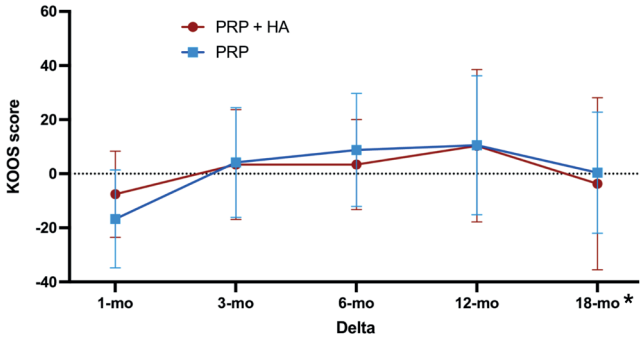

Fig. 4. The evolution of the mean Knee Injury and Osteoarthritis Outcome Score (KOOS) scores changes from baseline in the research groups

IKDC scale by timeframe

Two studies reported IKDC scores in the PRP + HA and PRP groups, and both studies were RCTs. Pooled analysis

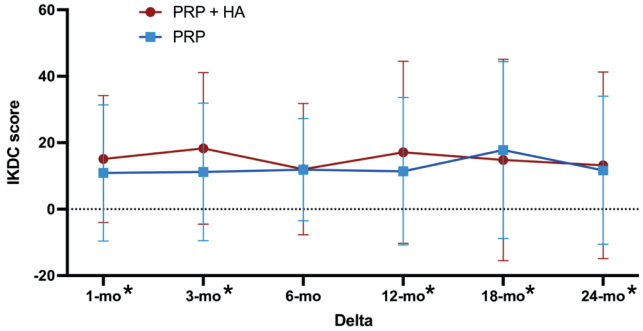

Fig. 5. The evolution of the mean International Knee Documentation Committee (IKDC) scores changes from baseline in the research groups

showed no significant difference between these groups at baseline (37.42 ±15.03 vs 37.47 ±10.42; SMD = 0.02; 95% CI: −0.35 to 0.39; p = 0.92), as well as across all follow-up periods (1, 3, 6, 12, 18, and 24 months; Supplementary Fig. 13). Changes in IKDC scores across different follow-up time points and baseline showed no statistically significant differences between the study groups in any of the outcome measures (Fig. 5).

Discussion

The presented meta-analysis offers an in-depth exploration into the effectiveness of PRP therapy, both as a standalone treatment and in combination with HA, in the management of KOA. This discussion integrates findings from the meta-analysis with a comprehensive review of existing literature to contextualize the impact of PRP and HA on patient outcomes, specifically focusing on pain reduction, functional improvement, and improvement in patients' overall quality of life.

The meta-analysis meticulously selected data from a broad spectrum of studies, narrowing down to 11 publications deemed suitable for rigorous scrutiny. This careful selection and synthesis process, which mirrored the comprehensive approach of Ciapini et al., not only underscores the current evidence but also reveals the nuanced benefits of PRP, with or without the addition of HA, in the management of KOA.[29]

When combined, PRP and HA jointly stimulate local angiogenesis, though PRP plays a more prominent role in this process by encouraging healthy surrounding cells to produce additional growth factors. Furthermore, PRP has demonstrated anti-nociceptive effects.[35–37] Research by Lana et al. suggests that HA serves as a supportive matrix for cartilage repair and improves the mechanical properties of cartilage.[32]

The observed improvements, as noted in the WOMAC Pain Assessment (WOMAC PA) scores, typically manifest between 30 and 90 days post-treatment, aligning with the timeframe for these molecular and cellular modifications to take effect. This combination therapy may result in better rehabilitation and earlier return to activities of daily living.

This meta-analysis unveils a notable variation in the efficacy of PRP across different timeframes and outcome measures. This observation resonates with the findings of Guo et al.,[30] who reported no significant disparities in functional outcomes between groups treated with PRP in combination with HA and those treated with PRP alone. However, they hinted at a potential for improved functional scores with the combination therapy. This trend suggests that while PRP alone can impart beneficial effects, its combination with HA might enhance or extend these outcomes, warranting further investigation, as suggested by Ciapini et al.[29] Furthermore, the meta-analysis highlights a critical period of enhanced efficacy specifically at the 6-month and 24-month follow-ups, particularly notable in PRP + HA treatments. This timeframe-dependent effectiveness aligns with observations made by Yu et al.[28] and Wu et al.,[3] who observed significant improvements in pain relief and functional scores at similar intervals. These findings suggest the potential for PRP and HA therapies to induce lasting therapeutic benefits, offering a robust alternative to the transient nature of symptomatic treatments currently used in KOA management paradigms.

Notably, the sub-analysis of RCTs within the meta-analysis revealed statistically significant differences in outcomes favoring PRP + HA combinations at specific follow-up intervals. This echoes the systematic approach of Lana et al.[32] and Jacob et al.,[31] which emphasized the value of RCTs in establishing the efficacy of PRP treatments. The consistent observation across these studies of improved outcomes with PRP + HA interventions reinforces the potential of combinatory therapies for improving KOA care.

The analysis extends to the nuanced roles of WOMAC and KOOS scores as outcome measures, reflecting patient-reported assessments of pain, stiffness, and functional capacity. The meta-analysis, through its elaborate examination of these scores, highlights the complex nature of KOA symptoms and how distinct treatments uniquely address these issues. Such an approach mirrors the detailed analyses conducted by Palco et al.[6] and Sun et al.,[34] which also utilized these scales to discern the therapeutic value of PRP and HA interventions. The consistency between these studies highlights the critical role of patient-centered outcomes in evaluating the success of KOA treatments.

Recent meta-analyses highlight the significant clinical utility of PRP and HA, underscoring the relevance of the topic explored in this meta-analysis. Bensa et al. demonstrated that corticosteroid injections for KOA yield results comparable to HA and PRP only in the short term, whereas PRP shows clear superiority in longer-term follow-ups.[38] Moreover, Qiao et al. found that PRP and HA, including when used in combination therapy, do not present a higher risk of treatment-related adverse events compared to placebo. It is important to note that the absence of elevated safety risks also applies to corticosteroid use.[39] Regarding the most effective intervention strategy, a meta-analysis by Tao et al. revealed that repeated PRP treatments are more effective than a single administration.[40]

Moreover, this study critically examines the risk of bias in the included studies, utilizing established tools such as the RoB2 and ROBINS-I. This thorough scrutiny of study quality ensures that conclusions drawn are based on evidence of the highest integrity, a principle that is paramount in the realm of evidence-based medicine. This commitment to maintaining strict methodological standards is reflected in the studies included in this review, all of which strive to contribute to a more reliable and comprehensive understanding of PRP and HA therapies in osteoarthritis management.

The assessment of pain and functional outcomes in most studies relied heavily on subjective scales such as the VAS and the WOMAC, which could potentially limit the objectivity of the results. The selection of the IKDC and WOMAC scores as endpoints (surrogate measures) was made due to their broad application in clinical studies in this field. The WOMAC score is a self-reported tool used by patients to evaluate their pain, stiffness, and physical function in relation to osteoarthritis. Patients rate each item based on their experience, typically using a Likert scale (0–4) or a VAS.[41] In contrast, the IKDC score combines physician assessments with patient-reported outcomes and assesses knee function, particularly after injuries such as ligament tears. This score integrates clinical evaluations with patients' subjective assessment of knee symptoms, functionality, and activity levels.[42] The use of objective assessment methods such as the single-leg stance (SLS) test, which was used in Sun et al.,[34] or the Biodex Balance System SD, used in Wu et al.,[3] may contribute to greater objectivity in research and a better comparison of both treatment methods. Furthermore, advanced imaging tests such as magnetic resonance imaging (MRI), despite their costs and time requirements, could also contribute to improving the quality of research. Adding a control group receiving saline injections, which is often used as a placebo in intra-articular research, could also set a clear baseline for judging how well the treatment works.

The selection of appropriate patients for the studies is critical. As indicated by Jacob et al., patient weight plays a crucial role due to its significant impact on articular cartilage damage.[31] Long-term studies should consider changes in patient body weight, as weight reduction alone can stop disease progression irrespective of the treatment. Palco et al. highlighted the importance of disease stage in analyzing treatment outcomes: most studies on the conservative treatment of KOA mainly include patients in the early stages, while stages III and IV are less represented. It is hypothesized that the use of combined PRP + HA may benefit individuals with advanced stages of the condition.[6] The selection and administration methods of treatments also present a significant consideration. The study by Wu et al. was probably the first study in which a different drug administration protocol was used, as PRP was administered 1 week prior to HA, diverging from the concurrent administration seen in prior studies.[3] This approach yielded comparable or superior results, highlighting the need for further optimization of administration protocols.

Future scientific research should focus, among other things, on deepening the understanding of the biological mechanisms underlying the effects of PRP and HA, both individually and in synergy. Gaining insight into these mechanisms may lead to the discovery of substances with similar potential applications in intra-articular injections. From a clinical standpoint, further investigation is needed into the optimal injection protocols, including timing intervals. For instance, it would be important to determine whether more intensive treatment could significantly reduce the risk of KOA progression, which holds critical clinical relevance. Additionally, there is a need for a more comprehensive analysis of the endpoints used in clinical trials. It seems essential to implement 2 types of measurement scales in such studies: one that is self-reported by patients, and another completed by clinicians, preferably independent raters not directly involved in the care of trial participants. Lastly, the use of radiologically based endpoints should be encouraged and combined with these measurement scales. Furthermore, the reliance on highly experienced physicians for performing high-frequency color Doppler imaging, without a standardized practitioner for all examinations, may introduce variability in the measurement outcomes. Coupled with the lack of long-term follow-up data to evaluate the enduring effects of the treatments and the possible unaccounted use of concurrent analgesic or anti-inflammatory medications, these factors support a cautious interpretation of the findings. They also highlight the need for further research with more rigorous methodologies and involving larger, more homogeneous study populations.

Limitations of the study

We acknowledge several limitations of this research that may impact the breadth of its applicability and the interpretation of its outcomes. First, the variability in pre-treatment inflammation levels and synovial conditions, potentially influenced by patients' NSAID or steroid use, introduces a degree of heterogeneity within the study cohorts, complicating direct comparisons. Additionally, the inclusion of patients receiving bilateral injections might increase population heterogeneity, though it reflects real-life clinical scenarios. The observed differences in the effectiveness of PRP combined with HA in treating different knee compartments, attributed to varying stress distributions, highlight the complexity of osteoarthritis treatment and underscore the necessity for individualized therapeutic approaches. Furthermore, some outcomes were reported by only 1 study; while these results were included for thoroughness, they do not constitute aggregated estimates. These instances were distinctly categorized and should be regarded as single-study results rather than meta-analytic compilations.

Conclusions

This meta-analysis underscores the efficacy of PRP, either independently or in conjunction with HA, as a viable treatment for KOA, especially in alleviating pain and enhancing functionality over time. Notable advantages were evident at 6 and 24 months, indicating a synergistic effect. Nonetheless, clinical implications must be evaluated with caution due to variability across the studies. To improve evidence-based treatment methods, future research should

focus on large, high-quality trials that investigate the underlying biological mechanisms of this therapy.

Use of AI and AI-assisted technologies

Not applicable.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.16979738. The package contains the following files:

Supplementary Fig. 1. A summary table of review authors' judgements for each risk of bias item for RCT.

Supplementary Fig. 2. A summary table of review authors' judgements for each risk of bias item for non-RCT.

Supplementary Fig. 3. A plot of the distribution of review authors' judgements across RCTs for each risk of bias item.

Supplementary Fig. 4. A plot of the distribution of review authors' judgements across non-RCTs for each risk of bias item.

Supplementary Fig. 5. Forest plot of VAS scores among PRP + HA and PRP patients (results derived from single studies do not constitute a pooled analysis and should be interpreted with caution due to the limited generalizability of such findings).

Supplementary Fig. 6. The evolution of the mean VAS scores changes from baseline in the research groups in RCTs sub-analysis.

Supplementary Fig. 7. Forest plot of WOMAC (total) scores among PRP + HA and PRP patients (results derived from single studies do not constitute a pooled analysis and should be interpreted with caution due to the limited generalizability of such findings).

Supplementary Fig. 8. The evolution of the mean WOMAC pains subscale changes from baseline in the research groups.

Supplementary Fig. 9. The evolution of the mean WOMAC stiffness subscale changes from baseline in the research groups.

Supplementary Fig. 10. The evolution of the mean WOMAC function subscale changes from baseline in the research groups.

Supplementary Fig. 11. Forest plot of KOOS scores among PRP + HA and PRP patients (results derived from single studies do not constitute a pooled analysis and should be interpreted with caution due to the limited generalizability of such findings).

Supplementary Fig. 12. The evolution of the mean KOOS scores changes from baseline in the research groups in RCTs sub-analysis.

Supplementary Fig. 13. Forest plot of IKDC scores among PRP + HA and PRP patients.

Supplementary Table 1. Pooled results of different scales among PRP + HA and PRP in subgroup of RCT trials.

ORCID iDs

Marcin Tomaszewski https://orcid.org/0009-0005-1011-6669
Jacek Smereka https://orcid.org/0000-0002-1427-4796
Mahdi Al-Jeabory https://orcid.org/0000-0003-4412-6409
Piotr Fudalej https://orcid.org/0009-0003-5361-9384
Michał Pruc https://orcid.org/0000-0002-2140-9732
Maciej W. Krupowies https://orcid.org/00009-0004-9774-0431
Zuzanna Gaca https://orcid.org/0009-0004-7200-0402
Łukasz Szarpak https://orcid.org/0000-0002-0973-5455

References

1. Abate M, Verna S, Schiavone C, Di Gregorio P, Salini V. Efficacy and safety profile of a compound composed of platelet-rich plasma and hyaluronic acid in the treatment for knee osteoarthritis (preliminary results). *Eur J Orthop Surg Traumatol*. 2015;25(8):1321–1326. doi:10.1007/s00590-015-1693-3
2. Dieppe PA, Lohmander LS. Pathogenesis and management of pain in osteoarthritis. *Lancet*. 2005;365(9463):965–973. doi:10.1016/S0140-6736(05)71086-2
3. Wu YT, Li TY, Lee KC, et al. Efficacy of a novel intra-articular administration of platelet-rich plasma one-week prior to hyaluronic acid versus platelet-rich plasma alone in knee osteoarthritis: A prospective, randomized, double-blind, controlled trial. *J Clin Med*. 2022;11(11):3241. doi:10.3390/jcm11113241
4. Alshami AM. Knee osteoarthritis related pain: A narrative review of diagnosis and treatment. *Int J Health Sci*. 2014;8(1):85–104. doi:10.12816/0006075
5. Prochazkova M, Zanvit P, Dolezal T, Prokesova L, Krsiak M. Increased gene expression and production of spinal cyclooxygenase 1 and 2 during experimental osteoarthritis pain. *Physiol Res*. 2009;58(3):419–425. doi:10.33549/physiolres.931439
6. Palco M, Fenga D, Basile GC, et al. Platelet-rich plasma combined with hyaluronic acid versus leucocyte and platelet-rich plasma in the conservative treatment of knee osteoarthritis: A retrospective study. *Medicina (Kaunas)*. 2021;57(3):232. doi:10.3390/medicina57030232
7. Robson EK, Hodder RK, Kamper SJ, et al. Effectiveness of weight-loss interventions for reducing pain and disability in people with common musculoskeletal disorders: A systematic review with meta-analysis. *J Orthop Sports Phys Ther*. 2020;50(6):319–333. doi:10.2519/jospt.2020.9041
8. Tschopp M, Pfirrmann CWA, Fucentese SF, et al. A randomized trial of intra-articular injection therapy for knee osteoarthritis. *Invest Radiol*. 2023;58(5):355–362. doi:10.1097/RLI.0000000000000942
9. Bahari Golamkaboudi A, Vojoudi E, Babaeian Roshani K, Porouhan P, Houshangi D, Barabadi Z. Current non-surgical curative regenerative therapies for knee osteoarthritis. *Stem Cell Rev Rep*. 2024;20(8):2104–2123. doi:10.1007/s12015-024-10768-6
10. Blaga FN, Nutiu AS, Lupsa AO, Ghiurau NA, Vlad SV, Ghitea TC. Exploring platelet-rich plasma therapy for knee osteoarthritis: An in-depth analysis. *J Funct Biomater*. 2024;15(8):221. doi:10.3390/jfb15080221
11. Soufan S, Al Khoury J, Hamdan Z, Rida MA. Intra-articular interventions in osteoarthritis: Navigating the landscape of hyaluronic acid, mesenchymal stem cells, and platelet-rich plasma. *World J Orthop*. 2024;15(8):704–712. doi:10.5312/wjo.v15.i8.704
12. Carr BJ, Miller AV, Colbath AC, Peralta S, Frye CW. Literature review details and supports the application of platelet-rich plasma products in canine medicine, particularly as an orthobiologic agent for osteoarthritis. *J Am Vet Med Assoc*. 2024;262(Suppl 1):S8–S15. doi:10.2460/javma.23.12.0692
13. Colen S, Van Den Bekerom MPJ, Mulier M, Haverkamp D. Hyaluronic acid in the treatment of knee osteoarthritis: A systematic review and meta-analysis with emphasis on the efficacy of different products. *BioDrugs*. 2012;26(4):257–268. doi:10.1007/BF03261884
14. Iturriaga V, Vásquez B, Bornhardt T, Del Sol M. Effects of low and high molecular weight hyaluronic acid on the osteoarthritic temporomandibular joint in rabbit. *Clin Oral Invest*. 2021;25(7):4507–4518. doi:10.1007/s00784-020-03763-x
15. Alves R, Grimalt R. A review of platelet-rich plasma: History, biology, mechanism of action, and classification. *Skin Appendage Disord*. 2018;4(1):18–24. doi:10.1159/000477353

16. Moussa M, Lajeunesse D, Hilal G, et al. Platelet rich plasma (PRP) induces chondroprotection via increasing autophagy, anti-inflammatory markers, and decreasing apoptosis in human osteoarthritic cartilage. *Exp Cell Res*. 2017;352(1):146–156. doi:10.1016/j.yexcr.2017.02.012
17. Muir SM, Reisbig N, Baria M, Kaeding C, Bertone AL. The concentration of plasma provides additional bioactive proteins in platelet and autologous protein solutions. *Am J Sports Med*. 2019;47(8):1955–1963. doi:10.1177/0363546519849671
18. Branch E, Emami A, Plummer H, Anz A. Paper 14: Leukocyte-poor platelet rich plasma versus leukocyte-poor platelet rich plasma plus hyaluronic acid for the treatment of symptomatic knee osteoarthritis: A prospective, randomized control trial with 2 year follow up. *Orthop J Sports Med*. 2022;10(7 Suppl 5):2325967121S00578. doi:10.1177/2325967121S00578
19. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71
20. Higgins JPT, Thomas J, Chandler J, et al., eds. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, UK: Wiley; 2019. doi:10.1002/9781119536604
21. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan: A web and mobile app for systematic reviews. *Syst Rev*. 2016;5(1):210. doi:10.1186/s13643-016-0384-4
22. Hefti F, Müller W, Jakob RP, Stäubli H -U. Evaluation of knee ligament injuries with the IKDC form. *Knee Surg Sports Traumatol Arthrosc*. 1993;1(3–4):226–234. doi:10.1007/BF01560215
23. Bellamy N, Wilson C, Hendrikz J. Population-Based Normative Values for the Western Ontario and McMaster (WOMAC) Osteoarthritis Index: Part I. *Semin Arthritis Rheum*. 2011;41(2):139–148. doi:10.1016/j.semarthrit.2011.03.002
24. Sterne JAC, Savović J, Page MJ, et al. RoB 2: A revised tool for assessing risk of bias in randomised trials. *BMJ*. 2019;366:l4898. doi:10.1136/bmj.l4898
25. Sterne JA, Hernán MA, Reeves BC, et al. ROBINS-I: A tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. doi:10.1136/bmj.i4919
26. Hozo SP, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. *BMC Med Res Methodol*. 2005;5(1):13. doi:10.1186/1471-2288-5-13
27. Higgins JPT, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. doi:10.1136/bmj.d5928
28. Yu W, Xu P, Huang G, Liu L. Clinical therapy of hyaluronic acid combined with platelet-rich plasma for the treatment of knee osteoarthritis. *Exp Ther Med*. 2018;16(3):2119–2125. doi:10.3892/etm.2018.6412
29. Ciapini G, Simonetti M, Giuntoli M, et al. Is the combination of platelet-rich plasma and hyaluronic acid the best injective treatment for grade II–III knee osteoarthritis? A prospective study. *Adv Orthop*. 2023;2023:1868943. doi:10.1155/2023/1868943
30. Guo Y, Yu H, Yuan L, et al. Treatment of knee osteoarthritis with platelet-rich plasma plus hyaluronic acid in comparison with platelet-rich plasma only. *Int J Clin Exp Med*. 2016;9(6):12085–12090. https://e-century.us/files/ijcem/9/6/ijcem0023424.pdf
31. Jacob G, Shetty V, Shetty S. A study assessing intra-articular PRP vs PRP with HMW HA vs PRP with LMW HA in early knee osteoarthritis. *J Arthrosc Joint Surg*. 2017;4(2):65–71. doi:10.1016/j.jajs.2017.08.008
32. Lana JFSD, Weglein A, Sampson SE, et al. Randomized controlled trial comparing hyaluronic acid, platelet-rich plasma and the combination of both in the treatment of mild and moderate osteoarthritis of the knee. *J Stem Cells Regen Med*. 2016;12(2):69–78. doi:10.46582/jsrm.1202011
33. Xu Z, He Z, Shu L, Li X, Ma M, Ye C. Intra-articular platelet-rich plasma combined with hyaluronic acid injection for knee osteoarthritis Is superior to platelet-rich plasma or hyaluronic acid alone in inhibiting inflammation and improving pain and function. *Arthroscopy*. 2021;37(3):903–915. doi:10.1016/j.arthro.2020.10.013
34. Sun SF, Lin GC, Hsu CW, Lin HS, Liou IHS, Wu SY. Comparing efficacy of intraarticular single crosslinked Hyaluronan (HYAJOINT Plus) and platelet-rich plasma (PRP) versus PRP alone for treating knee osteoarthritis. *Sci Rep*. 2021;11(1):140. doi:10.1038/s41598-020-80333-x
35. Murali A, Khan I, Tiwari S. Navigating the treatment landscape: Choosing between platelet-rich plasma (PRP) and hyaluronic acid (HA) for knee osteoarthritis management. A narrative review. *J Orthop Rep*. 2024;3(1):100248. doi:10.1016/j.jorep.2023.100248
36. Bennell KL, Hunter DJ, Paterson KL. Platelet-rich plasma for the management of hip and knee osteoarthritis. *Curr Rheumatol Rep*. 2017;19(5):24. doi:10.1007/s11926-017-0652-x
37. Say F, Gürler D, Yener K, Bülbül M, Malkoc M. Platelet-rich plasma injection is more effective than hyaluronic acid in the treatment of knee osteoarthritis. *Acta Chir Orthop Traumatol Cech*. 2013;80(4):278–283. PMID:24119476.
38. Bensa A, Sangiorgio A, Boffa A, Salerno M, Moraca G, Filardo G. Corticosteroid injections for knee osteoarthritis offer clinical benefits similar to hyaluronic acid and lower than platelet-rich plasma: A systematic review and meta-analysis. *EFORT Open Rev*. 2024;9(9):883–895. doi:10.1530/EOR-23-0198
39. Qiao X, Yan L, Feng Y, et al. Efficacy and safety of corticosteroids, hyaluronic acid, and PRP and combination therapy for knee osteoarthritis: A systematic review and network meta-analysis. *BMC Musculoskelet Disord*. 2023;24(1):926. doi:10.1186/s12891-023-06925-6
40. Tao X, Aw AAL, Leeu JJ, Bin Abd Razak HR. Three doses of platelet-rich plasma therapy are more effective than one dose of platelet-rich plasma in the treatment of knee osteoarthritis: A systematic review and meta-analysis. *Arthroscopy*. 2023;39(12):2568–2576.e2. doi:10.1016/j.arthro.2023.05.018
41. Bellamy N, Buchanan WW, Goldsmith CH, Campbell J, Stitt LW. Validation study of WOMAC: A health status instrument for measuring clinically important patient relevant outcomes to antirheumatic drug therapy in patients with osteoarthritis of the hip or knee. *J Rheumatol*. 1988;15(12):1833–1840. PMID:3068365.
42. Irrgang JJ, Anderson AF, Boland AL, et al. Development and validation of the International Knee Documentation Committee Subjective Knee Form. *Am J Sports Med*. 2001;29(5):600–613. doi:10.1177/03635465010290051301

Arthroscopic vs open surgery for shoulder dislocation and instability: A network meta-analysis of treatment outcomes

Chenghong Wen[A,D–F], Qiang Hua[B,C,F], Wenduo Qian[B,C,F], Jide Su[B,C,F], Mingming Lei[A,D–F]

Department of Sports Injury, Affiliated Sport Hospital of Chengdu Sport 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(5):765–778

Address for correspondence
Chenghong Wen
E-mail: wenchenghong610@hotmail.com

Funding sources
None declared

Conflict of interest
None declared

Received on July 5, 2024
Reviewed on December 1, 2024
Accepted on July 24, 2025

Published online on April 7, 2026

Abstract

Background. No comprehensive comparative research has been conducted to evaluate open Bankart (OB), arthroscopic Bankart (AB), open Latarjet (OL), and arthroscopic Latarjet (AL) simultaneously across all relevant clinical outcomes and parameters.

Objectives. To compare the efficacy of OB, AB, OL, and AL procedures in the treatment of shoulder dislocation.

Materials and methods. The databases PubMed, Embase, the Cochrane Library, and Web of Science were utilized for the literature search. The study evaluated recurrent instability, re-dislocation, apprehension, functional outcomes, and postoperative pain. The results were visually represented through network diagrams, forest plots, league tables, and rank probability plots to provide a comprehensive understanding of each outcome.

Results. Overall, 37 studies were included in the analysis. Individuals who underwent OL experienced a notably reduced risk of recurrent instability compared with those who underwent AB (random-effects model pooled relative risk (RR) = 0.34, 95% credible interval (95% CrI): 0.24–0.48) and OB (random-effects model pooled RR = 0.51, 95% CrI: 0.31–0.85). The risk of re-dislocation was also significantly lower for patients treated with OL compared with AB (pooled RR = 0.15, 95% CrI: 0.04–0.45). While not statistically significant, the OL procedure tended to have the lowest risk of apprehension and the highest Subjective Shoulder Value (SSV) score. Regarding postoperative pain, patients who underwent OB had the highest likelihood of attaining the lowest scores on the visual analogue scale (VAS). In addition, OL was associated with the highest probability of complications.

Conclusions. The open Latarjet procedure appears to offer superior shoulder stability; however, while functional outcomes for patients undergoing OL are likely to be non-inferior, the procedure is not significantly associated with reduced postoperative pain as measured with the VAS score. Additionally, the OL procedure is associated with an increased likelihood of complications. Consequently, it is essential to implement preventive measures to manage postoperative pain and address potential complications following OL procedure.

Key words: network meta-analysis, shoulder dislocation, Bankart, Open Latarjet, arthroscopic treatment

Cite as
Wen C, Hua Q, Qian W, Su J, Lei M. Arthroscopic vs open surgery for shoulder dislocation and instability: A network meta-analysis of treatment outcomes. *Adv Clin Exp Med*. 2026;35(5):765–778. doi:10.17219/acem/208614

DOI
10.17219/acem/208614

Highlights

- Open Latarjet procedure provides superior shoulder stability with significantly lower rates of recurrent instability and re-dislocation compared to Bankart techniques.
- Arthroscopic and open Bankart repairs show higher recurrence risk, while open Latarjet demonstrates the most favorable stability outcomes in shoulder dislocation treatment.
- Postoperative pain outcomes vary, with open Bankart associated with lower pain scores, while open Latarjet does not significantly reduce pain despite better stability.
- Open Latarjet carries a higher risk of complications, highlighting the need for careful patient selection and post-operative management strategies.

Introduction

The shoulder joint is the most frequently dislocated joint in the human body, with an incidence of approx. 23.96 per 100,000 individuals annually.[1] Its high mobility and relatively shallow glenoid cavity make it particularly susceptible to instability, especially following traumatic events.[1] Injuries to the static (e.g., labrum, ligaments) and dynamic (e.g., muscles) stabilizers of the shoulder can lead to instability, with anterior shoulder instability being the most common form, typically resulting from traumatic events.[2,3] Conservative management is frequently selected for patients who are not candidates for surgery, despite the higher risk of recurrence.[4] Surgical stabilization represents an effective treatment option when conservative management fails, providing durable fixation and a more definitive resolution of instability.[5]

Traditional open Bankart (OB) repair was historically regarded as the standard surgical treatment for shoulder stabilization.[6,7] This technique has been shown to improve joint stability and is associated with low recurrence rates.[8–11] However, despite its effectiveness, the open approach has certain disadvantages, including restricted external rotation and an increased risk of secondary osteoarthritis.[12]

Arthroscopic Bankart (AB) repair, first described in 1993, has gained increasing acceptance over the past decades owing to advances in arthroscopic techniques and improved surgical expertise.[13–15] Compared with open procedures, arthroscopic repair offers several advantages, including smaller incisions, shorter operative time, reduced postoperative pain, and potentially fewer complications.[16–18] However, some studies have reported higher recurrence rates of shoulder instability following AB repair compared with open repair.[19,20]

The open Latarjet (OL) procedure is widely used for the management of anterior shoulder instability and is recognized for its effectiveness and reliability; however, it is associated with potential complications such as nonunion and infection.[21,22] Arthroscopic Latarjet (AL) has subsequently been introduced as a minimally invasive alternative. Emerging evidence indicates that AL achieves clinical outcomes comparable to OL, while potentially offering smaller incisions, fewer complications, faster graft healing, earlier rehabilitation, and the ability to address concomitant intra-articular pathologies.[23,24]

Numerous studies have directly compared pairs of the 4 surgical techniques – OB, AB, OL, and AL – for the management of shoulder instability.[25–27] Furthermore, several meta-analyses have evaluated the comparative effectiveness of OL compared with AL in recurrent anterior shoulder instability, as well as OB vs AB in the management of Bankart lesions.[28,29]

Currently, multiple surgical interventions are available for managing shoulder dislocations; however, no single treatment has been proven superior. Additionally, there is a scarcity of studies evaluating OB, AB, OL, and AL across all relevant clinical outcomes and parameters.

Objectives

The current network meta-analysis aimed to compare the OB, AB, OL, and AL procedures for the treatment of shoulder dislocation.

Materials and methods

Search strategy

To ensure a thorough examination of the available literature, 2 independent investigators (Q.H. and W.Q.) conducted a comprehensive search across multiple databases, including PubMed, Embase, the Cochrane Library, and Web of Science. The search was completed on August 17, 2023, to gather the most recent and relevant studies for the network meta-analysis. English search terms included "Shoulder Dislocation" OR "Dislocation, Shoulder" OR "Dislocations, Shoulder" OR "Shoulder Dislocations" OR "Glenohumeral Dislocation" OR "Dislocation, Glenohumeral" OR "Dislocations, Glenohumeral" OR "Glenohumeral Dislocations" OR "Glenohumeral Subluxation" OR "Glenohumeral Subluxations" OR "Subluxation, Glenohumeral" OR "Subluxations, Glenohumeral" AND "Bristow" OR "Latarjet" OR "Bankart" OR "iliac bone graft"

OR "Repair" OR "Putti-Platt" OR "Arthroscopies" OR "arthroscopic surgery" OR "Surgery" OR "Conservative" OR "nonoperative" OR "nonsurgical" OR "Physiotherapy" OR "Immobilization". Disagreements concerning eligibility were resolved by another investigator (J.S.). Primary study selection was based on the titles and abstracts of the retrieved studies, followed by full-text screening. The present study adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) reporting guidelines, ensuring a transparent and methodologically rigorous approach to the systematic review and meta-analysis process.[30]

Eligibility criteria

The inclusion criteria for the study were as follows: 1) studies involving patients diagnosed with shoulder instability or dislocation; 2) studies in which patients underwent treatment with one of the 4 specified surgical procedures: OB, AB, OL, or AL; 3) evaluation of outcomes related to shoulder stability, functional outcomes, or postoperative pain; 4) study designs limited to randomized controlled trials (RCTs) and cohort studies.

The exclusion criteria were defined to ensure the relevance and quality of the included literature and included: 1) animal studies; 2) case reports; 3) studies with sample sizes of fewer than 10 participants; 4) studies not directly related to the topic, such as those involving patients without shoulder dislocation or instability, or those focusing on non-surgical treatments or procedures other than Bankart and Latarjet repairs; 5) editorial materials, conference abstracts, protocols, letters, guidelines, expert consensus documents, reviews, and meta-analyses, as these formats typically do not contain the level of detail or original data required for a systematic review and meta-analysis; 6) studies not published in English.

Outcome measures

In this analysis, the primary outcome of interest was shoulder stability, with specific measures including the incidence of recurrent instability, re-dislocation, and apprehension. The secondary outcomes encompassed functional outcomes and postoperative pain. Functional outcomes were measured using several standardized scales, including the Subjective Shoulder Value (SSV) score, reflecting the patient's assessment of shoulder function; the American Shoulder and Elbow Surgeons (ASES) score, a comprehensive evaluation of shoulder function commonly used after shoulder surgery; the Rowe score, another scale specifically designed to assess shoulder function, with an emphasis on activities of daily living and sports; and forward flexion, a measure of range of motion indicating the extent to which the patient can raise the arm forward. Postoperative pain was quantified using the visual analogue scale (VAS).

Data extraction

The data extraction process from the included studies in the systematic review and network meta-analysis was comprehensive and included the following details: authors, year of publication, country, study design, population, intervention, sample size, male/female distribution, age (in years), body mass index (BMI, in kg/m^2), glenoid bone loss, number of Hill–Sachs lesions, follow-up time (in months), and outcomes.

Quality assessment

In the systematic review and network meta-analysis, the quality assessment of the included RCTs was conducted using the modified Jadad scale.[31] This scale allocates a total of 7 points, with RCT quality categorized as follows: 1–3 points indicating low quality and 4–7 points indicating high quality. For cohort studies, quality was evaluated using the modified Newcastle–Ottawa scale (NOS).[32] The NOS assigns a total of 9 points, with quality ratings defined as follows: 0–3 points indicating poor quality, 4–6 points indicating fair quality, and 7–9 points indicating good quality.

Statistical analyses

A network meta-analysis was conducted using Stata v. 15.1 (StataCorp LLC, College Station, USA) and the gemtc 1.0.1 package in R v. 4.1.3 (R Foundation for Statistical Computing, Vienna, Austria). A Bayesian framework and a Markov chain Monte Carlo (MCMC) model were employed for the analysis. To ensure model convergence, trace plots and diagnostic tools were utilized. The number of initial iterations in the Monte Carlo simulations was determined based on achieving a potential scale reduction factor (PSRF) close to 1, indicating satisfactory convergence. Additional iterations were performed if necessary.

The bandwidth value in the density plots served as a quantitative measure, with smaller values suggesting a closer fit to the preset distribution. Iterations were continued until the bandwidth stabilized near 0, ensuring robust model performance. The model consisted of 4 chains, with an initial iteration count of 20,000, followed by an additional 50,000 iterations, and a step length of 1. In the presence of network relationships, the model assessed consistency and potential discrepancies between direct and indirect treatment comparisons. When the difference between the Deviance Information Criteria (DIC) of the consistency and inconsistency models was less than 5, it indicated agreement between direct and indirect comparisons. Network plots were constructed for each outcome measure. Analyses were conducted using a random-effects model. For the Rowe score, SSV score, forward flexion, ASES score, and VAS, the weighted

mean difference (WMD) and 95% credible intervals (95% CrIs) for different surgical interventions were reported. For recurrent instability, apprehension, re-dislocation, and complications, the relative risk (RR) values and 95% CrIs for different surgical interventions were reported. All direct and indirect comparisons of WMD or RR values with their corresponding 95% CrIs were presented in forest plots. The ranking of surgical interventions was predicted using a ranking plot, in which the probability of each intervention being ranked in the nth position was presented as a bar chart, with the horizontal axis indicating the corresponding rank.

Continuous variables were analyzed using the change from baseline as the final analysis outcome to reduce statistical errors arising from initial population imbalances in some cohort studies. An adjusted funnel plot was employed to evaluate potential publication bias. This approach is particularly relevant in the context of network meta-analyses, where traditional funnel plots may not adequately capture the nuances of multiple intervention comparisons. By using the adjusted funnel plot, symmetry of effect sizes across studies can be assessed more accurately, thereby providing a more robust evaluation of potential bias.

Results

Characteristics of the included studies

From the 4 databases, 8,402 studies were retrieved. After excluding duplicates, 5,982 studies remained for screening based on titles and abstracts, followed by full-text screening. Ultimately, 37 studies were included in this network meta-analysis, comprising 5 RCTs and 32 cohort studies.[25–27,33–66] Figure 1 illustrates the flow chart of the eligible study screening process. Of the included studies, 32 were of high quality and 5 were of fair quality. These studies were published between 1997 and 2023. The characteristics of the included studies are presented in Supplementary Table 1.

Shoulder stability

Recurrent instability

Information on recurrent instability was provided in 21 studies involving 2,300 patients, in which OB, AB, OL, and AL were compared. Arthroscopic Bankart was directly compared with OB, AL, and OL. There were more studies for the direct comparison between AB and OL (Fig. 2A). The forest plot analysis using a random-effects

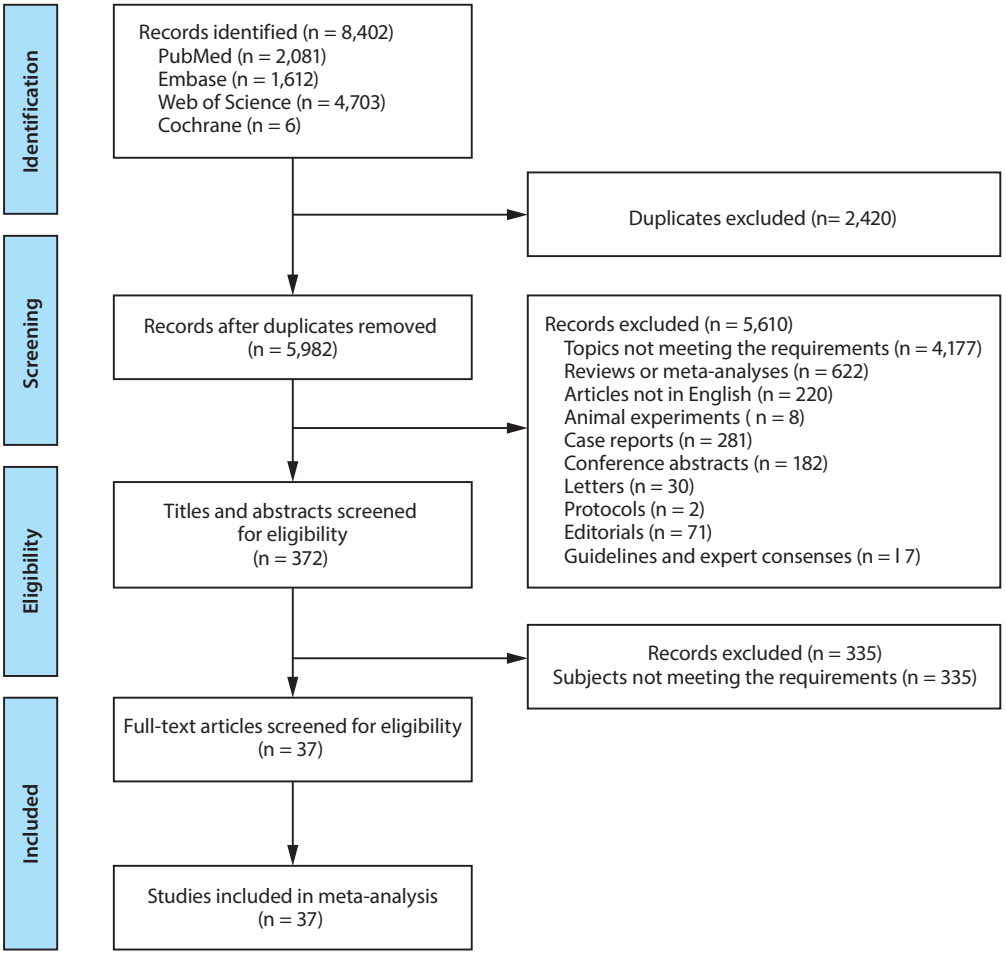

Fig. 1. Flowchart of eligible study screening

Table 1. League table of different procedures for various outcomes analyzed by the random-effect model

Outcomes/procedures		AB	AL	OB	OL
Recurrent instability	AB	AB	0.72 (0.22, 2.52)	0.67 (0.42, 1.04)	0.34 (0.24, 0.48)
	AL	1.38 (0.38, 4.57)	AL	0.94 (0.24, 3.3)	0.47 (0.13, 1.57)
	OB	1.47 (0.95, 2.36)	1.07 (0.3, 4.18)	OB	0.51 (0.31, 0.85)
	OL	2.95 (2.07, 4.18)	2.13 (0.64, 7.77)	1.99 (1.17, 3.34)	OL
Apprehension	AB	AB	0.76 (0.13, 3.4)	1.88 (0.24, 19.33)	0.57 (0.18, 1.45)
	AL	1.31 (0.29, 7.75)	AL	2.49 (0.21, 50.16)	0.75 (0.22, 2.68)
	OB	0.53 (0.05, 4.16)	0.4 (0.02, 4.83)	OB	0.3 (0.02, 2.79)
	OL	1.75 (0.69, 5.64)	1.33 (0.37, 4.56)	3.33 (0.36, 48.3)	OL
Re-dislocation	AB	AB	0.25 (0, 28.83)	0.65 (0.18, 2.09)	0.15 (0.04, 0.45)
	AL	3.96 (0.03, 524.99)	AL	2.55 (0.02, 360.41)	0.59 (0.01, 62.89)
	OB	1.55 (0.48, 5.55)	0.39 (0, 51.8)	OB	0.23 (0.04, 1.04)
	OL	6.71 (2.25, 27.26)	1.73 (0.02, 183.9)	4.34 (0.96, 24.12)	OL
Complications	AB	AB	0.98 (0.27, 4.77)	1.27 (0.39, 4.19)	2.08 (0.97, 6.26)
	AL	1.02 (0.21, 3.71)	AL	1.31 (0.18, 6.78)	2.12 (0.77, 6.53)
	OB	0.79 (0.22, 2.67)	0.76 (0.15, 5.41)	OB	1.64 (0.48, 8.38)
	OL	0.48 (0.15, 1.01)	0.47 (0.14, 1.32)	0.61 (0.12, 2.13)	OL
Rowe scores	AB	AB	−2.36 (−12.83, 8.38)	5.92 (−5.56, 16.63)	3.77 (−2.49, 10.28)
	AL	2.36 (−8.58, 12.83)	AL	8.3 (−7.65, 23.12)	6.22 (−3.79, 15.77)
	OB	−5.92 (−16.63, 5.46)	−8.3 (−23.12, 7.65)	OB	−2.14 (−14.5, 11.14)
	OL	−3.77 (−10.28, 2.49)	−6.12 (−15.77, 3.79)	2.14 (−11.14, 14.5)	OL
SSV score [%]	AB	AB	3.83 (−32.47, 40.42)	6.97 (−13.12, 27.37)	–
	AL	−3.83 (−40.42, 32.47)	AL	3.1 (−27.24, 33.4)	–
	OL	−6.97 (−27.37, 13.12)	−3.1 (−33.4, 27.24)	OL	–
Forward flexion	AB	AB	−14.36 (−32.77, 3.36)	−1.68 (−12.05, 7.64)	–
	AL	14.36 (−3.36, 32.77)	AL	12.6 (−2.54, 27.61)	–
	OL	1.68 (−7.65, 12.05)	−12.6 (−27.61, 2.58)	OL	–
ASES score	AB	AB	−22.88 (−47.99, 2.49)	−13.8 (−31.3, 3.96)	–
	AL	22.88 (−2.49, 47.99)	AL	9.1 (−8.8, 27.07)	–
	OL	13.8 (−3.96, 31.3)	−9.1 (−27.07, 8.8)	OL	–
VAS	AB	AB	−0.40 (−3.24, 2.44)	0.29 (−1.74, 2.31)	–
	OB	0.4 (−2.44, 3.24)	OB	0.69 (−2.81, 4.19)	–
	OL	−0.29 (−2.31, 1.75)	−0.69 (−4.19, 2.81)	OL	–

SSV – subjective shoulder value; ASES – American Shoulder and Elbow Surgeons; VAS – visual analogue scale; OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet. Values in brackets are 95% credible intervals (95 CrIs).

model revealed that OL had a lower risk of recurrent instability than AB (pooled relative risk (RR) = 0.33, 95% CrI: 0.22–0.49) (Fig. 3). The league table demonstrated that patients treated with OL had a significantly lower risk of recurrent instability than those treated with AB (random-effects model pooled RR = 0.34, 95% CrI: 0.24–0.48) and OB (random-effects model pooled RR = 0.51, 95% CrI: 0.31–0.85) (Table 1). The rank probabilities showed that OL was most likely to be the optimal procedure with regard to recurrent instability (Table 2). In the subgroup analysis restricted to primary surgical procedures, the pooled results analyzed using a random-effects model demonstrated that patients undergoing OL exhibited a significantly reduced risk of recurrent instability compared with those treated with arthroscopic AB (pooled RR = 0.39, 95% CrI: 0.20–0.60) (Supplementary Table 2).

Re-dislocation

Nineteen studies involving 2,318 patients assessed the risk of re-dislocation after AB, OB, AL, and OL. Arthroscopic Bankart was directly compared with OB and OL, and OL was directly compared with OB and AL. More studies compared AB and OL (Fig. 2B). Figure 4, based on the random-effects model analysis, also showed a lower risk of re-dislocation in patients treated with OL compared with AB (pooled RR = 0.13, 95% CrI: 0.03–0.43). The league table

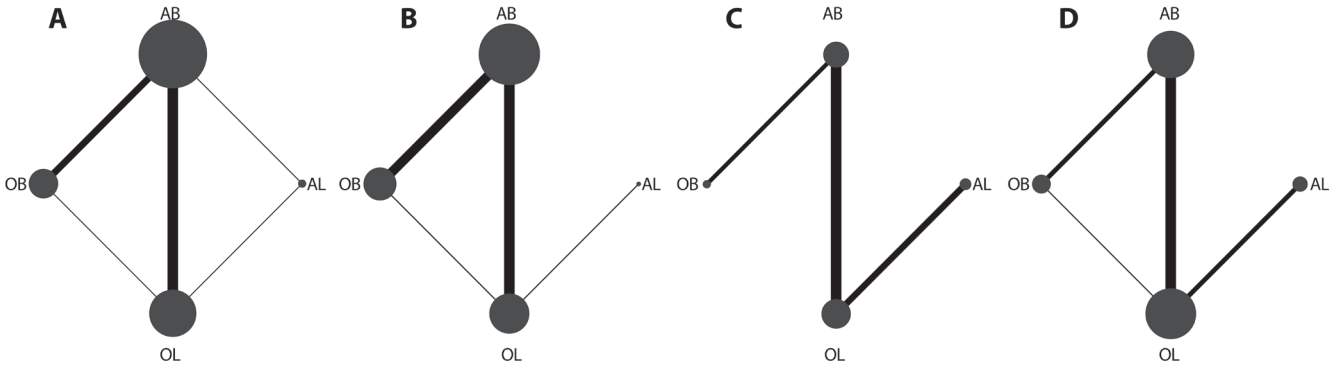

Fig. 2. Network plots of different procedures for shoulder stability. A. recurrent instability; B. re-dislocation; C. apprehension; D. complications

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet.

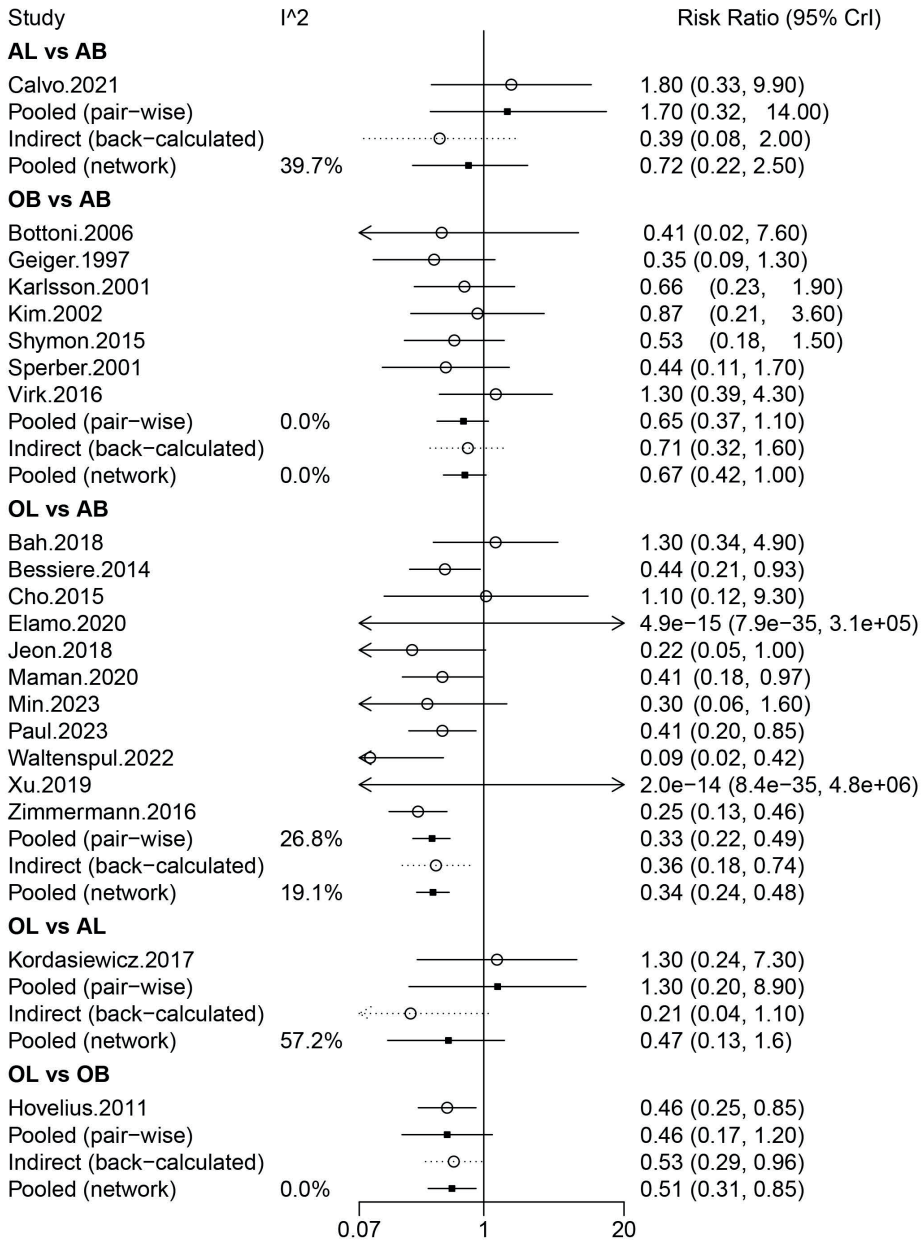

Fig. 3. Forest plots of different procedures for shoulder stability analyzed with the random-effect model analysis (recurrent instability)

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet; CrI – credibility interval.

analysis using a random-effects model found that the risk of re-dislocation after OL was significantly lower than that after AB (pooled RR = 0.15, 95% CrI: 0.04–0.45) (Table 1). The rank probabilities showed that OL had the highest likelihood of being the optimal operation with regard to re-dislocation (Table 2). Based on the subgroup analysis

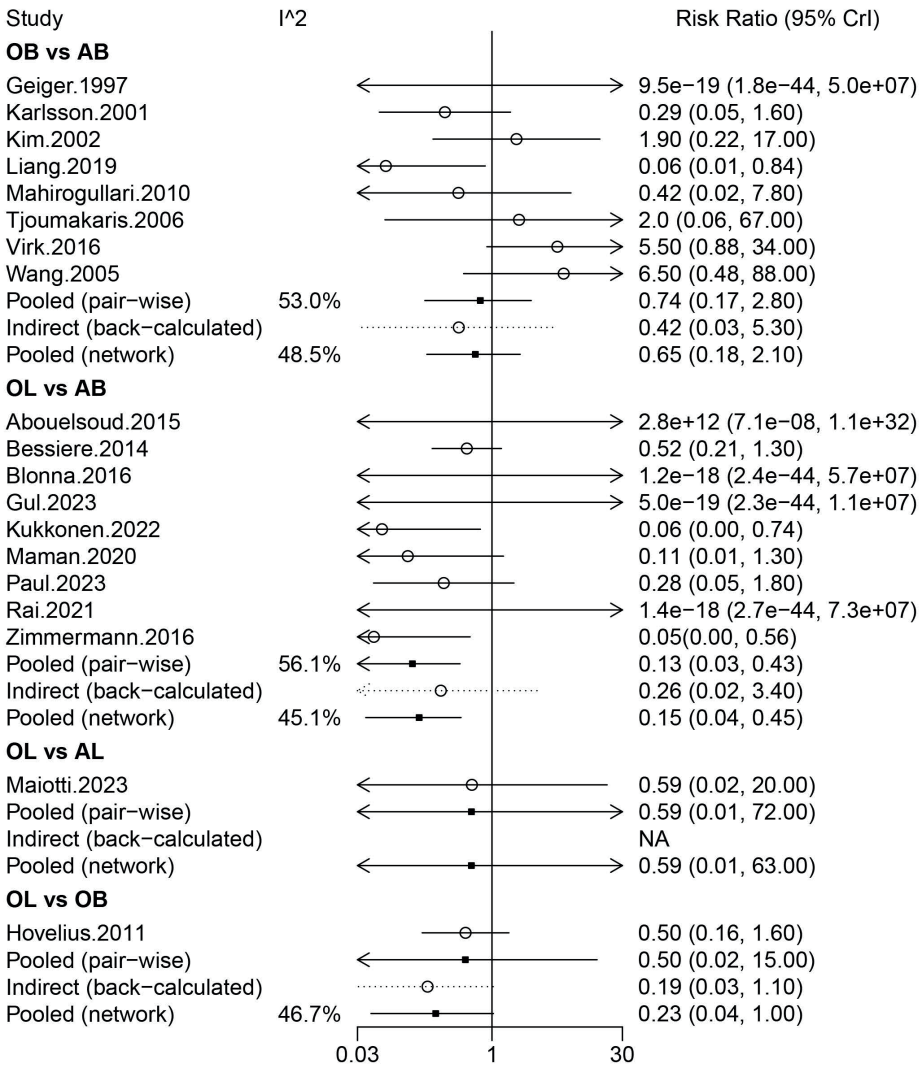

Fig. 4. Forest plots of different procedures for shoulder stability analyzed with the random-effect model analysis (re-dislocation)

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet; CrI – credibility interval.

of primary surgeries, OL was associated with a lower likelihood of re-dislocation compared with AB (random-effects model pooled RR = 0.14, 95% CrI: 0.03–0.39) and OB (random-effects model pooled RR = 0.11, 95% CrI: 0.01–0.66) (Supplementary Table 2).

Apprehension

The risk of apprehension was evaluated in 10 studies including 1,195 patients. Arthroscopic Bankart was directly compared with OL and OB, and AL was directly compared with OL. More studies compared AB and OL (Fig. 2C). The forest plot analysis using a random-effects model demonstrated that no significant difference was observed in the risk of apprehension among these 4 operations (Fig. 5). Similarly, no significant difference was observed in the risk of apprehension among these 4 operations in the league table analysis using a random-effects model (Table 1). The rank probabilities showed that patients undergoing OL were most likely to have the lowest risk of apprehension (Table 2).

Functional outcomes

SSV score

Three studies involving 500 patients investigated the SSV score, including AB, OL, and AL. No significant difference in the SSV score was found among these surgical treatments based on the forest plot and league table (Table 1). The rank probabilities suggested that OL was most likely to be associated with the highest SSV score (Table 2). Based on the subgroup analysis of primary surgeries, no significant difference in the SSV score was observed among these surgical treatments (Supplementary Table 2).

ASES score

Two studies involving 952 patients assessed the ASES score, including AB, OL, and AL. No significant difference was observed in the ASES score among these surgical treatments based on the league table analysis using

Table 2. Rank probabilities of different procedures for various outcomes analyzed using the random-effect model

Outcomes/procedures		[1]	[2]	[3]	[4]
Recurrent instability	AB	0.67388	0.313325	0.012795	0
	AL	0.295145	0.24836	0.35018	0.106315
	OB	0.03096	0.43655	0.52601	0.00648
	OL	0.000015	0.001765	0.111015	0.887205
Apprehension	AB	0.167745	0.54864	0.226875	0.05674
	AL	0.14133	0.24713	0.35775	0.25379
	OB	0.679765	0.132705	0.07371	0.11382
	OL	0.01116	0.071525	0.341665	0.57565
Re-dislocation	AB	0.579025	0.359475	0.06113	0.00037
	AL	0.25713	0.090375	0.25431	0.398185
	OB	0.163505	0.53854	0.280285	0.01767
	OL	0.00034	0.01161	0.404275	0.583775
Complications	AB	0.01209	0.18025	0.46991	0.33775
	AL	0.05388	0.24929	0.248675	0.448155
	OB	0.20487	0.328335	0.253815	0.21298
	OL	0.72916	0.242125	0.0276	0.001115
Rowe scores	AB	0.01475	0.15198	0.56608	0.26719
	AL	0.046865	0.10945	0.201175	0.64251
	OB	0.62036	0.20063	0.10547	0.07354
	OL	0.318025	0.53794	0.127275	0.01676
SSV score [%]	AB	0.160565	0.313975	0.52546	–
	AL	0.38385	0.23618	0.37997	
	OL	0.455585	0.449845	0.09457	
Forward flexion	AB	0.70733	0.25892	0.03375	––
	AL	0.025975	0.035085	0.93894	
	OL	0.266695	0.705995	0.02731	
ASES score	AB	0.934345	0.04722	0.018435	–
	AL	0.02623	0.0955	0.87827	
	OL	0.039425	0.85728	0.103295	
VAS	AB	0.235695	0.52824	0.236065	–
	OB	0.263845	0.18881	0.547345	
	OL	0.50046	0.28295	0.21659	

SSV – subjective shoulder value; ASES – American Shoulder and Elbow Surgeons; VAS – visual analogue scale; OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet.
[1] – the probability that the procedure ranks 1st (i.e., performs the best) for that outcome; [2] – the probability that the procedure ranks 2nd for that outcome; [3] – the probability that the procedure ranks 3rd for that outcome; [4] – the probability that the procedure ranks 4th (i.e., performs the worst) for that outcome.

a random-effects model (Table 1). The rank probabilities suggested that AB was most likely to be the best surgery with regard to the ASES score (Table 2).

Rowe score

The Rowe score was evaluated in 11 studies involving 1,152 patients, and comparisons were made among OB, AB, OL, and AL. No significant difference in the Rowe score was found among these operations when the analysis was conducted using a random-effects model (Fig. 6). Similarly, no significant difference in the Rowe score was observed among these operations in the league table analysis using a random-effects model (Table 1). The rank probabilities showed that OB had the highest likelihood of being the optimal procedure with regard to the Rowe score (Table 2).

Forward flexion

Three studies involving 311 patients provided data on forward flexion, including AB, OL, and AL. No significant difference in forward flexion was observed among the 3 operations in the forest plot analysis using

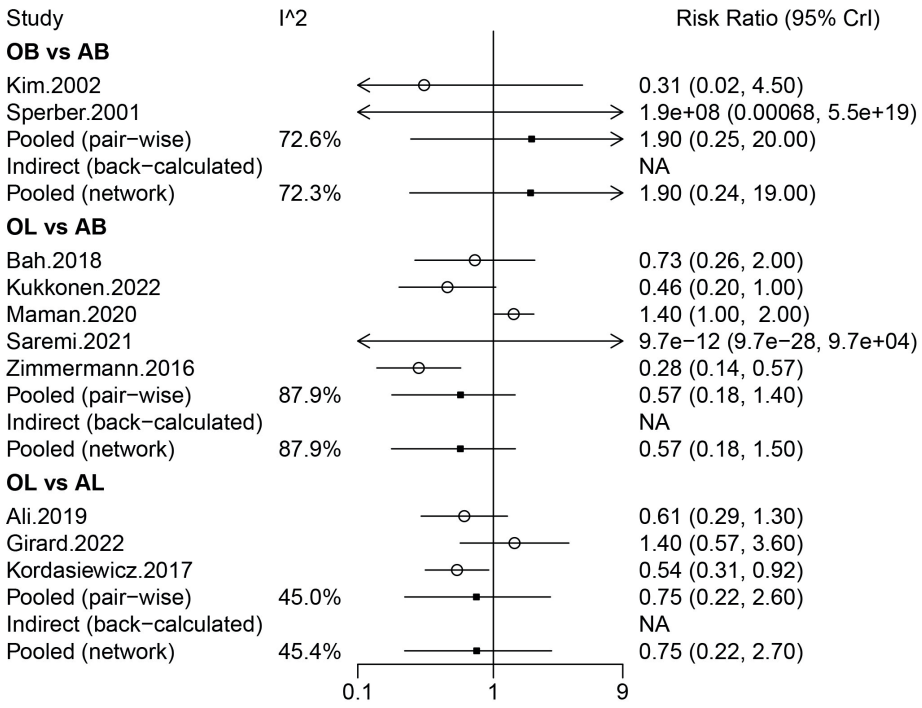

Fig. 5. Forest plots of different procedures for shoulder stability analyzed with the random-effect model analysis (apprehension)

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet; CrI – credibility interval.

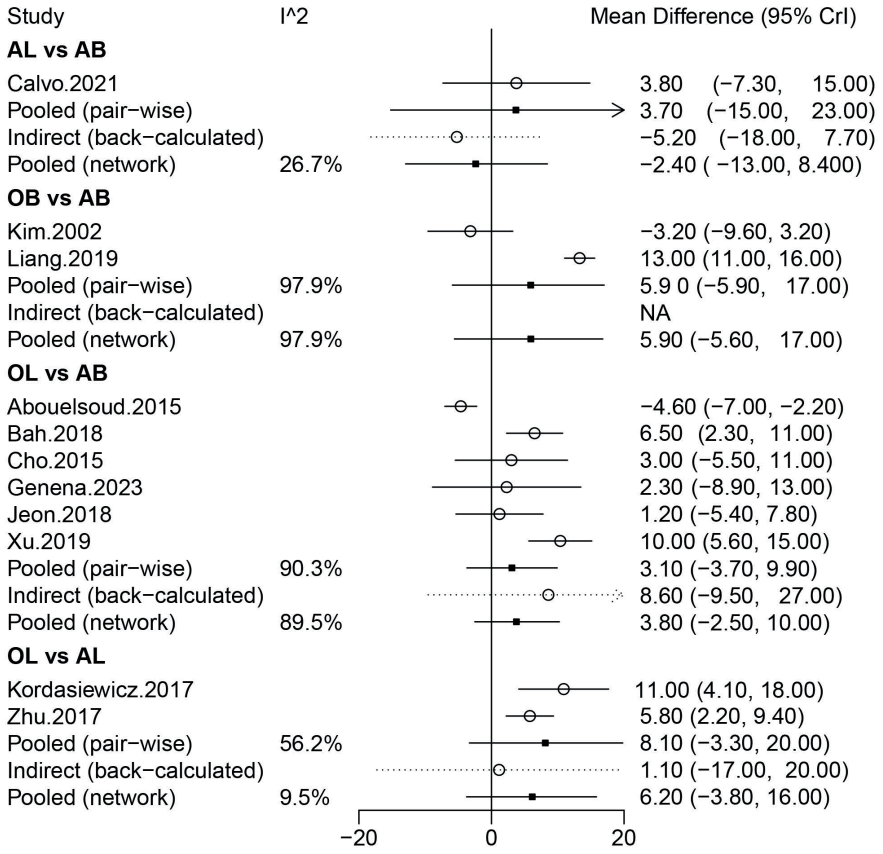

Fig. 6. Forest plots of different procedures for shoulder stability analyzed with the random-effect model analysis (Rowe score)

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet; CrI – credibility interval.

a random-effects model (Fig. 7). Similarly, no significant difference in forward flexion was observed among the 3 operations in the league table analysis using a random-effects model (Table 1). The rank probabilities showed that AL was most likely to be associated with the greatest forward flexion (Table 2).

Postoperative pain

Information on the VAS score was provided in 3 studies involving 1,294 patients, including AB, OL, and OB. When the analysis was conducted using a random-effects model, no significant difference was observed in the VAS score

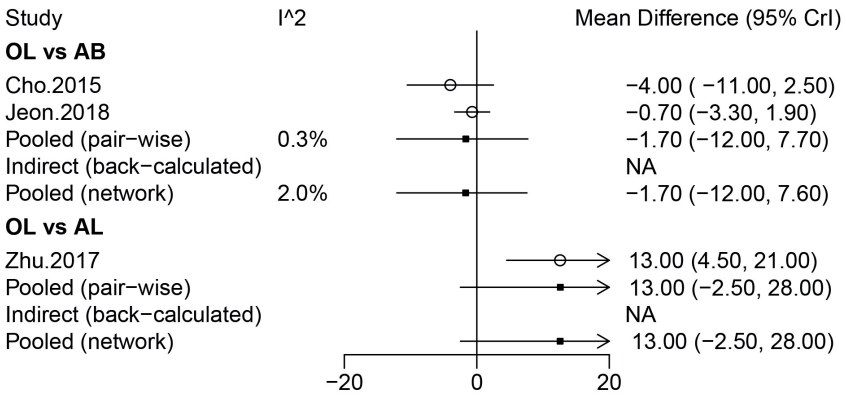

Fig. 7. Forest plots of different procedures for shoulder stability analyzed with the random-effect model analysis (forward flexion)

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet; CrI – credibility interval.

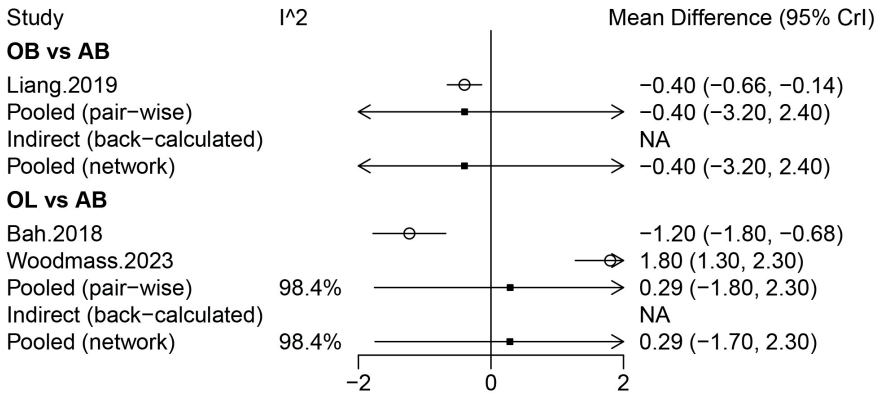

Fig. 8. Forest plots of different procedures for shoulder stability analyzed with the random-effect model analysis (visual analogue scale (VAS) score)

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet; CrI – credibility interval.

among AB, OL, and OB (Fig. 8). Similarly, no significant difference in the VAS score was observed among AB, OL, and OB in the league table analysis using a random-effects model (Table 1). The rank probabilities demonstrated that OB had the highest probability of achieving the lowest VAS score (Table 2).

Complications

A total of 18 studies comprising 1,902 patients were included in the analysis. Complications were assessed across 4 surgical approaches: AB, OB, OL, and AL. Direct comparisons were made between OL and each of the other 3 surgical methods, with specific direct comparisons existing between OL and AB. The network plot indicated a larger node and a thicker connecting line between AB and OL, suggesting that more studies directly compared these 2 interventions and that the sample size for these comparisons was relatively larger (Fig. 2D).

No significant difference in complications was observed among the 4 methods in the forest plot analysis using a random-effects model (Fig. 9). Similarly, no significant difference in complications was observed among the 4 techniques in the league table analysis using a random-effects model (Table 1). The rank probabilities demonstrated that OL had the highest probability of complications (Table 2). The funnel plots for publication bias assessment are shown in Supplementary Fig. 1.

Discussion

In this network meta-analysis, a comparison of the pooled data regarding the clinical efficacy of the OB, AB, OL, and AL procedures for shoulder dislocation and instability was conducted. Open Latarjet was most likely to have the lowest risk of recurrent instability and re-dislocation. Although not statistically significant, OL was most likely to have the lowest risk of apprehension and the highest SSV score. However, OL had the highest probability of complications. In addition, with regard to postoperative pain, the OL procedure was not necessarily associated with a significantly lower VAS score, indicating that it might not be the best option in terms of postoperative pain management.

The results of the network meta-analysis indicate that the OL procedure may offer superior outcomes in terms of shoulder stability and function for patients with shoulder dislocation and instability. However, the analysis also highlights a critical area for improvement: the management of postoperative pain following the OL procedure. In a meta-analysis conducted by Wang et al.,[29] a comparative analysis of OB compared with AB repairs for Bankart lesions revealed that while the OB method provided superior shoulder stability, it was accompanied by limitations in shoulder mobility.

Conversely, the AB technique, although less invasive, was associated with better preservation of motion. In another

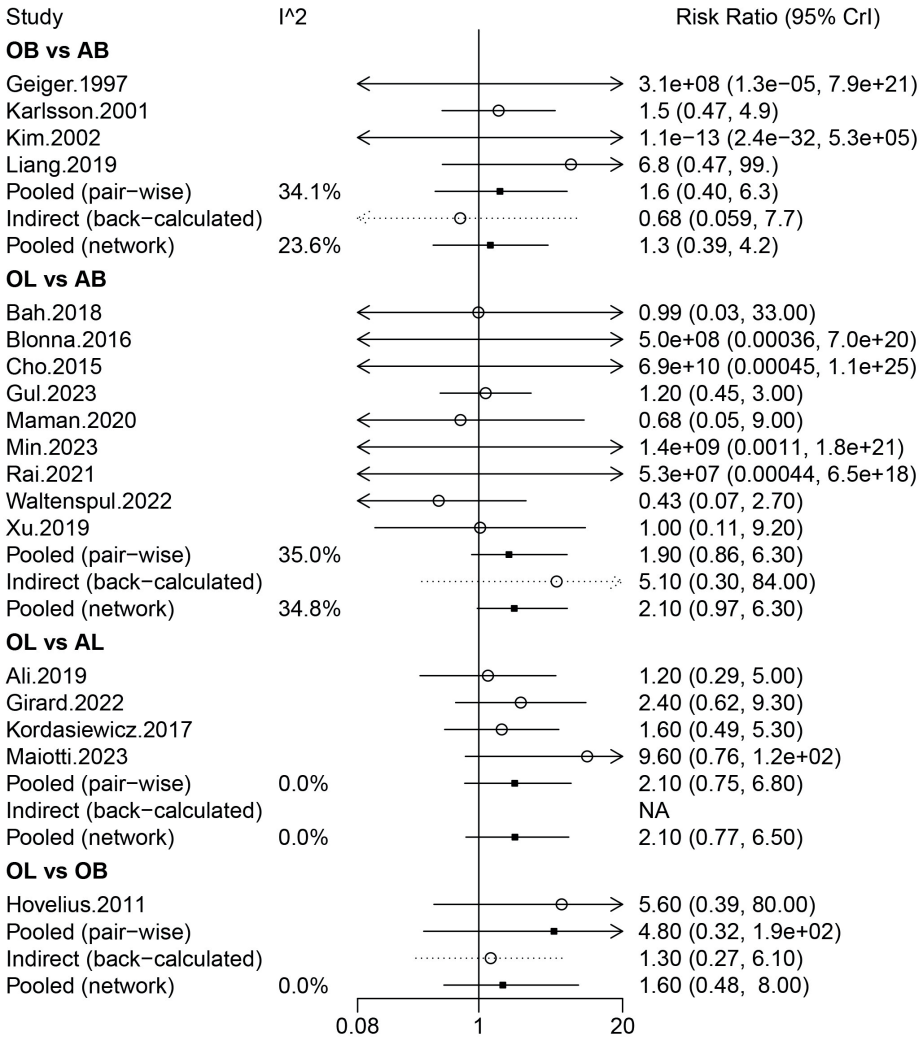

Fig. 9. Forest plots of different procedures for shoulder stability analyzed with the random-effect model analysis (complications)

OB – open Bankart; AB – arthroscopic Bankart; OL – open Latarjet; AL – arthroscopic Latarjet; CrI – credibility interval.

meta-analysis, a comprehensive evaluation was conducted to compare the clinical efficacy, as well as the rates of postoperative revisions and complications, between OL and AL. The study found no significant differences in the majority of the assessed outcomes between the 2 procedures. However, the AL approach was associated with a lower score on the Western Ontario Shoulder Instability Index, indicating better shoulder function or less instability. Additionally, the AL procedure exhibited a higher revision rate, suggesting a greater likelihood of patients requiring additional surgeries.[28]

Nonetheless, the efficacy of OB, AB, OL, and AL has not been compared in patients with shoulder dislocation and instability in existing studies. To facilitate understanding and decision-making among the 4 methods when considering therapeutic options for shoulder dislocation, the present network meta-analysis focused on shoulder stability, functional outcomes, and postoperative pain following these procedures.

In terms of shoulder stability, the OL procedure was indicated to have the highest likelihood of being the most effective repair option. In a meta-analysis encompassing 795 shoulders, the OL procedure demonstrated a significantly lower risk of recurrence compared with AB repair.[67] As reported by Rollick et al.,[68] the estimated re-dislocation rate was notably higher at 15.1% for patients who had undergone AB, whereas it was considerably lower at 2.7% for those who had OL procedure. In terms of functional outcomes, OL may also demonstrate a relative advantage.

According to the findings by Bliven et al.,[69] a higher percentage of patients who underwent the Latarjet procedure were able to successfully return to work, sports, and throwing activities compared with those who had Bankart repair. Consistently, a study has shown that the Latarjet procedure for anterior shoulder instability leads to excellent long-term functional outcomes and a high rate of return to sports.[70] However, patients might experience unsatisfactory postoperative pain outcomes and complications following the OL procedure, which could be attributed to the larger incision associated with the surgery. Arthroscopic surgery generally offers the benefits of a lower nonunion rate and the potential for quicker recovery due to its minimally invasive nature. The drawbacks of the AL procedure are also evident. It requires longer operative time and a longer learning curve for surgeons.[71]

Analysis of the learning curve indicated that the initial cohort experienced longer operative times and higher complication rates compared with the subsequent group,

potentially accounting for the increased incidence of revision surgeries required for the AL procedure.[72] Only surgeons expected to perform the AL procedure at high volume should consider adopting this technique.[73,74] Additionally, it has been shown that the direct costs associated with the AL procedure are significantly higher, amounting to approximately double the cost of the OL procedure, with figures of 2,335 EUR for AL compared with 1,040 EUR for OL.[75]

Several factors may contribute to the effectiveness of the OL procedure. The OL procedure provides a dual stabilizing effect by increasing the anterior glenoid width and creating a sling effect with the conjoint tendon, which helps stabilize the joint.[76] Correct graft positioning is critical to the success of the OL procedure. It has been suggested that OL might offer better control over graft placement, which could lead to improved long-term outcomes.[77]

Limitations of the study

Several limitations of this study should be acknowledged. First, there was a lack of literature on some outcomes, which may have affected the accuracy of the results. Second, this study focused on the general population, and further research was not conducted in specific populations such as athletes and males or females.[51] In addition, some of the included studies were cohort studies; therefore, change scores were used for continuous data, and studies that provided only endpoint data were not included in this analysis. Third, including only English-language literature may result in language bias. Fourth, there are differences between primary operations and re-operations, and these differences could potentially influence the outcomes of interest. However, due to limitations in the included literature, such as the lack of detailed data on re-operations, we were unable to conduct a separate analysis for re-operations. This precludes definitive conclusions regarding the comparative effectiveness of the 4 surgical approaches for shoulder dislocation and instability in the context of re-operations. Fifth, we acknowledge that glenoid bone loss, a significant factor affecting shoulder joint stability and surgical outcomes, was not controlled for in our screening process. Although glenoid bone loss was included as a variable in Supplementary Table 1, the substantial amount of missing data across the reviewed studies prevented us from accounting for it in our analysis. This limitation may restrict the accuracy of our comparisons among different surgical approaches.

Conclusions

Open Latarjet was most likely to be the optimal procedure for shoulder stability, and patients undergoing OL may have non-inferior functional outcomes. Although the OL procedure may not be associated with significantly lower VAS pain scores, it was associated with a higher rate of complications. Therefore, it is imperative to implement appropriate measures to mitigate postoperative pain and manage complications following OL procedure.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.15872256. The package contains the following files:

Supplementary Table 1. Characteristics of the included studies.

Supplementary Table 2. League table of different procedures for various outcomes in patients with primary operation analyzed by the random-effect model.

Supplementary Fig. 1. A. Publication bias assessed by the funnel plot for recurrent instability; B. Publication bias assessed by the funnel plot for re-dislocation; C. Publication bias assessed by the funnel plot for apprehension; D. Publication bias assessed by the funnel plot for Rowe; E. Publication bias assessed by the funnel plot for complications.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Chenghong Wen https://orcid.org/0009-0006-7136-3302
Wenduo Qian https://orcid.org/0009-0002-7371-4154
Mingming Lei https://orcid.org/0009-0006-4088-4236
Qiang Hua https://orcid.org/0000-0002-4582-0382
Jide Su https://orcid.org/0009-0007-5512-1837

References

1. Patrick CM, Snowden J, Eckhoff MD, et al. Epidemiology of shoulder dislocations presenting to United States emergency departments: An updated ten-year study. *World J Orthop*. 2023;14(9):690–697. doi:10.5312/wjo.v14.i9.690
2. Stokes DJ, McCarthy TP, Frank RM. Physical therapy for the treatment of shoulder instability. *Phys Med Rehabil Clin North Am*. 2023;34(2):393–408. doi:10.1016/j.pmr.2022.12.006
3. Christopher HW, Grainger AJ. Anterior shoulder instability. *Semin Musculoskelet Radiol*. 2022;26(5):546–557. doi:10.1055/s-0042-1756168
4. Do W, Kim J, Lim J, Yoon T, Shin S, Chun Y. High failure rate after conservative treatment for recurrent shoulder dislocation without subjective apprehension on physical examination. *Knee Surg Sports Traumatol Arthrosc*. 2023;31(1):178–184. doi:10.1007/s00167-022-07028-w
5. Belk JW, Wharton BR, Houck DA, et al. Shoulder stabilization versus immobilization for first-time anterior shoulder dislocation: A systematic review and meta-analysis of level 1 randomized controlled trials. *Am J Sports Med*. 2023;51(6):1634–1643. doi:10.1177/03635465211065403
6. Fares MY, Boufadel P, Daher M, Koa J, Khanna A, Abboud JA. Anterior shoulder instability and open procedures: History, indications, and clinical outcomes. *Clin Orthop Surg*. 2023;15(4):521. doi:10.4055/cios23018
7. AlSomali K, Kholinne E, Van Nguyen T, et al. Outcomes and return to sport and work after open Bankart repair for recurrent shoulder instability: A systematic review. *Orthop J Sports Med*. 2021;9(10):23259671211026907. doi:10.1177/23259671211026907
8. Monk AP, Crua E, Gatenby GC, et al. Clinical outcomes following open anterior shoulder stabilization for glenohumeral instability in the young collision athlete. *J Shoulder Elbow Surg*. 2022;31(7):1474–1478. doi:10.1016/j.jse.2021.12.013

9. Mancini MR, Arciero RA. Open Bankart repair. *Clin Sports Med*. 2024; 43(4):617–633. doi:10.1016/j.csm.2023.12.002
10. Bitar I, Bustos D, Marangoni L, Robles C, Gentile L, Bertiche P. Outcomes of open Bankart repair plus inferior capsular shift compared with Latarjet procedure in contact athletes with recurrent anterior shoulder instability. *Arch Bone Joint Surg*. 2023;11(1):39–46. doi:10.22038/abjs.2022.60208.2974
11. Neviaser AS, Benke MT, Neviaser RJ. Open Bankart repair for revision of failed prior stabilization: Outcome analysis at a mean of more than 10 years. *J Shoulder Elbow Surg*. 2015;24(6):897–901. doi:10.1016/j.jse.2014.11.036
12. Berendes T, Mathijssen N, Verburg H, Kraan G. The open-modified Bankart procedure: Long-term follow-up 'a 16–26-year follow-up study.' *Arch Orthop Trauma Surg*. 2018;138(5):597–603. doi:10.1007/s00402-017-2866-9
13. Hu B, Hong J, Zhu H, Yan S, Wu H. Arthroscopic Bankart repair versus conservative treatment for first-time traumatic anterior shoulder dislocation: A systematic review and meta-analysis. *Eur J Med Res*. 2023;28(1):260. doi:10.1186/s40001-023-01160-0
14. Rashid MS, Arner JW, Millett PJ, Sugaya H, Emery R. The Bankart repair: Past, present, and future. *J Shoulder Elbow Surg*. 2020;29(12):e491–e498. doi:10.1016/j.jse.2020.06.012
15. Vermeulen AE, Landman EBM, Veen EJD, Nienhuis S, Koorevaar CT. Long-term clinical outcome of arthroscopic Bankart repair with suture anchors. *J Shoulder Elbow Surg*. 2019;28(5):e137–e143. doi:10.1016/j.jse.2018.09.027
16. McQuivey KS, Brinkman JC, Tummala SV, Shaha JS, Tokish JM. Arthroscopic remplissage using knotless, all-suture anchors. *Arthrosc Techn*. 2022;11(4):e615–e621. doi:10.1016/j.eats.2021.12.015
17. Su BY, Yi SY, Peng T, Yi G, Zhang L. Comparison of arthroscopic surgery versus open surgical repair of the anterior talofibular ligament: A retrospective study of 80 patients from a single center. *Med Sci Monit*. 2020;27:e928526. doi:10.12659/msm.928526
18. Ahmed AF, Alzobi OZ, Hantouly AT, et al. Complications of elbow arthroscopic surgery: A systematic review and meta-analysis. *Orthop J Sports Med*. 2022;10(11):23259671221137863. doi:10.1177/23259671221137863
19. Hurley ET, Manjunath AK, Bloom DA, et al. Arthroscopic Bankart repair versus conservative management for first-time traumatic anterior shoulder instability: A systematic review and meta-analysis. *Arthroscopy*. 2020;36(9):2526–2532. doi:10.1016/j.arthro.2020.04.046
20. Haskel JD, Wang KH, Hurley ET, et al. Clinical outcomes of revision arthroscopic Bankart repair for anterior shoulder instability: A systematic review of studies. *J Shoulder Elbow Surg*. 2022;31(1):209–216. doi:10.1016/j.jse.2021.06.021
21. Werthel JD, Sabatier V, Schoch B, et al. Outcomes of the Latarjet procedure for the treatment of chronic anterior shoulder instability: Patients with prior arthroscopic Bankart repair versus primary cases. *Am J Sports Med*. 2020;48(1):27–32. doi:10.1177/0363546519888909
22. Longo UG, Loppini M, Rizzello G, Ciuffreda M, Maffulli N, Denaro V. Latarjet, Bristow, and Eden–Hybinette procedures for anterior shoulder dislocation: Systematic review and quantitative synthesis of the literature. *Arthroscopy*. 2014;30(9):1184–1211. doi:10.1016/j.arthro.2014.04.005
23. Cerciello S, Corona K, Morris BJ, Santagada DA, Maccauro G. Early outcomes and perioperative complications of the arthroscopic Latarjet procedure: Systematic review and meta-analysis. *Am J Sports Med*. 2019;47(9):2232–2241. doi:10.1177/0363546518783743
24. Griesser MJ, Harris JD, McCoy BW, et al. Complications and re-operations after Bristow–Latarjet shoulder stabilization: A systematic review. *J Shoulder Elbow Surg*. 2013;22(2):286–292. doi:10.1016/j.jse.2012.09.009
25. Maiotti M, De Vita A, De Benedetto M, et al. Clinical outcomes and recurrence rate of 4 procedures for recurrent anterior shoulder instability: ASA, remplissage, open, and arthroscopic Latarjet. A multicenter study. *J Shoulder Elbow Surg*. 2023;32(5):931–938. doi:10.1016/j.jse.2022.10.030
26. Saremi H, Saneii A, Goodarzi B. Midterm clinical results of Bankart repair, Bankart remplissage, and Latarjet procedures for treating recurrent anterior shoulder dislocation. *Adv Hum Biol*. 2021;11(Suppl 1):S22–S26. doi:10.4103/aihb.aihb_22_21
27. Woodmass JM, Wagner ER, Smith J, et al. Postoperative recovery comparisons of arthroscopic Bankart to open Latarjet for the treatment of anterior glenohumeral instability. *Eur J Orthop Surg Traumatol*. 2022;33(4):1357–1364. doi:10.1007/s00590-022-03265-4
28. Deng Z, Zheng Y, Su J, et al. Open versus arthroscopic Latarjet for recurrent anterior shoulder instability: A systematic review and meta-analysis. *Orthop J Sports Med*. 2023;11(5):23259671231174476. doi:10.1177/23259671231174476
29. Wang L, Liu Y, Su X, Liu S. A meta-analysis of arthroscopic versus open repair for treatment of Bankart lesions in the shoulder. *Med Sci Monit*. 2015;21:3028–3035. doi:10.12659/msm.894346
30. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*. 2021;372:n71. doi:10.1136/bmj.n71
31. Jadad AR, Moore RA, Carroll D, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Cont Clin Trials*. 1996;17(1):1–12. doi:10.1016/0197-2456(95)00134-4
32. Wells GA, Shea BJ, O'Connell D, Peterson J, Tugwell P. The Newcastle–Ottawa Scale (NOS) for Assessing the Quality of Non-Randomized Studies in Meta-Analysis. Ottawa, Canada: The Ottawa Hospital Research Institute; 2020. https://ohri.ca/en/who-we-are/core-facilities-and-platforms/ottawa-methods-centre/newcastle-ottawa-scale. Accessed March 6, 2026.
33. Paul RW, Reddy MP, Sonnier JH, et al. Increased rates of subjective shoulder instability after Bankart repair with remplissage compared to Latarjet surgery. *J Shoulder Elbow Surg*. 2023;32(5):939–946. doi:10.1016/j.jse.2022.11.001
34. Min KS, Wake J, Cruz C, et al. Surgical treatment of shoulder instability in active-duty service members with subcritical glenoid bone loss: Bankart vs Latarjet. *J Shoulder Elbow Surg*. 2023;32(4):771–775. doi:10.1016/j.jse.2022.10.011
35. Gul Y, Farooq MZ,. I, Essa MA, Khan SU. Evaluation of arthroscopic Bankart repair and open Latarjet technique for treatment of recurrent shoulder dislocation. *Pak J Med Health Sci*. 2023;17(2):528–530. doi:10.53350/pjmhs2023172528
36. Genena A, Hashem M, Waly A, Hegazy MO. Open Latarjet versus arthroscopic Bankart repair for the treatment of traumatic anterior shoulder instability in high-demand patients with minimal glenoid bone loss. *Cureus*. 2023;15(4):e37127. doi:10.7759/cureus.37127
37. Waltenspül M, Ernstbrunner L, Ackermann J, Thiel K, Galvin JW, Wieser K. Long-term results and failure analysis of the open Latarjet procedure and arthroscopic Bankart repair in adolescents. *J Bone Joint Surg*. 2022;104(12):1046–1054. doi:10.2106/jbjs.21.01050
38. Kukkonen J, Elamo S, Flinkkilä T, et al. Arthroscopic Bankart versus open Latarjet as a primary operative treatment for traumatic anteroinferior instability in young males: A randomised controlled trial with 2-year follow-up. *Br J Sports Med*. 2022;56(6):327–333. doi:10.1136/bjsports-2021-104028
39. Girard M, Dalmas Y, Martinel V, Mansat P, Bonnevialle N. Arthroscopic Latarjet with cortical buttons versus open Latarjet with screws: A short-term comparative study. *Am J Sports Med*. 2022;50(12):3326–3332. doi:10.1177/03635465221120076
40. Rai S, Tamang N, Sharma LK, et al. Comparative study of arthroscopic Bankart repair versus open Latarjet procedure for recurrent shoulder dislocation. *J Int Med Res*. 2021;49(4):3000605211007328. doi:10.1177/03000605211007328
41. Calvo E, Luengo G, Morcillo D, Foruria AM, Valencia M. Revision arthroscopic bankart repair versus arthroscopic Latarjet for failed primary arthroscopic stabilization with subcritical bone loss. *Orthop J Sports Med*. 2021;9(5):23259671211001809. doi:10.1177/23259671211001809
42. Maman E, Dolkart O, Krespi R, et al. A multicenter retrospective study with a minimum 5-year follow-up comparing arthroscopic Bankart repair and the Latarjet procedure. *Orthop J Sports Med*. 2020;8(8):2325967120941366. doi:10.1177/2325967120941366
43. Elamo S, Selänne L, Lehtimäki K, et al. Bankart versus Latarjet operation as a revision procedure after a failed arthroscopic Bankart repair. *JSES Int*. 2020;4(2):292–296. doi:10.1016/j.jseint.2020.01.004
44. Ali J, Altintas B, Pulatkan A, Boykin RE, Aksoy DO, Bilsel K. Open versus arthroscopic Latarjet procedure for the treatment of chronic anterior glenohumeral instability with glenoid bone loss. *Arthroscopy*. 2020;36(4):940–949. doi:10.1016/j.arthro.2019.09.042

45. Xu Y, Wu K, Ma Q, et al. Comparison of clinical and patient-reported outcomes of three procedures for recurrent anterior shoulder instability: Arthroscopic Bankart repair, capsular shift, and open Latarjet. *J Orthop Surg Res*. 2019;14(1):326. doi:10.1186/s13018-019-1340-5
46. Liang Z, Li B, Muheremu A. Arthroscopic versus open bankart repair in patients with recurrent anterior dislocation of the shoulder: A ten year follow-up study. *Int J Clin Exp Med*. 2019;12(6):7814–7819. https://e-century.us/files/ijcem/12/6/ijcem0089952.pdf.
47. Jeon YS, Jeong HY, Lee DK, Rhee YG. Borderline glenoid bone defect in anterior shoulder instability: Latarjet procedure versus Bankart repair. *Am J Sports Med*. 2018;46(9):2170–2176. doi:10.1177/0363546518776978
48. Bah A, Lateur GM, Kouevidjin BT, et al. Chronic anterior shoulder instability with significant Hill–Sachs lesion: Arthroscopic Bankart with remplissage versus open Latarjet procedure. *Orthop Traumatol Surg Res*. 2018;104(1):17–22. doi:10.1016/j.otsr.2017.11.009
49. Zhu Y, Jiang C, Song G. Arthroscopic versus open Latarjet in the treatment of recurrent anterior shoulder dislocation with marked glenoid bone loss: A prospective comparative study. *Am J Sports Med*. 2017;45(7):1645–1653. doi:10.1177/0363546517693845
50. Kordasiewicz B, Małachowski K, Kicinski M, Chaberek S, Pomianowski S. Comparative study of open and arthroscopic coracoid transfer for shoulder anterior instability (Latarjet): Clinical results at short term follow-up. *Int Orthop*. 2017;41(5):1023–1033. doi:10.1007/s00264-016-3372-3
51. Zimmermann SM, Scheyerer MJ, Farshad M, Catanzaro S, Rahm S, Gerber C. Long-term restoration of anterior shoulder stability: A retrospective analysis of arthroscopic Bankart repair versus open Latarjet procedure. *J Bone Joint Surg Am*. 2016;98(23):1954–1961. doi:10.2106/JBJS.15.01398
52. Virk MS, Manzo RL, Cote M, et al. Comparison of time to recurrence of instability after open and arthroscopic Bankart repair techniques. *Orthop J Sports Med*. 2016;4(6):2325967116654114. doi:10.1177/2325967116654114
53. Cho NS, Yoo JH, Rhee YG. Management of an engaging Hill–Sachs lesion: Arthroscopic remplissage with Bankart repair versus Latarjet procedure. *Knee Surg Sports Traumatol Arthrosc*. 2016;24(12):3793–3800. doi:10.1007/s00167-015-3666-9
54. Blonna D, Bellato E, Caranzano F, Assom M, Rossi R, Castoldi F. Arthroscopic Bankart repair versus open Bristow–Latarjet for shoulder instability: A matched-pair multicenter study focused on return to sport. *Am J Sports Med*. 2016;44(12):3198–3205. doi:10.1177/0363546516658037
55. Shymon SJ, Roocroft J, Edmonds EW. Traumatic anterior instability of the pediatric shoulder: A comparison of arthroscopic and open bankart repairs. *J Pediatr Orthop*. 2015;35(1):1–6. doi:10.1097/BPO.0000000000000215
56. Abouelsoud MM, Abdelrahman AA. Recurrent anterior shoulder dislocation with engaging Hill–Sachs defect: Remplissage or Latarjet? *Eur Orthop Traumatol*. 2015;6(3):151–156. doi:10.1007/s12570-015-0313-3
57. Bessière C, Trojani C, Carles M, Mehta SS, Boileau P. The open Latarjet procedure is more reliable in terms of shoulder stability than arthroscopic Bankart repair. *Clin Orthop Relat Res*. 2014;472(8):2345–2351. doi:10.1007/s11999-014-3550-9
58. Hovelius L, Vikerfors O, Olofsson A, Svensson O, Rahme H. Bristow-Latarjet and Bankart: A comparative study of shoulder stabilization in 185 shoulders during a seventeen-year follow-up. *J Shoulder Elbow Surg*. 2011;20(7):1095–1101. doi:10.1016/j.jse.2011.02.005
59. Mahiroğulları M, Ozkan H, Akyüz M, Uğraş AA, Güney A, Kuşkucu M. Comparison between the results of open and arthroscopic repair of isolated traumatic anterior instability of the shoulder. *Acta Orthop Traumatol Turc*. 2010;44(3):180–185. doi:10.3944/AOTT.2010.2289
60. Tjoumakaris FP, Abboud JA, Hasan SA, Ramsey ML, Williams GR. Arthroscopic and open Bankart repairs provide similar outcomes. *Clin Orthop Relat Res*. 2006;446:227–232. doi:10.1097/01.blo.0000205883.73705.19
61. Bottoni CR, Smith EL, Berkowitz MJ, Towle RB, Moore JH. Arthroscopic versus open shoulder stabilization for recurrent anterior instability: A prospective randomized clinical trial. *Am J Sports Med*. 2006;34(11):1730–1737. doi:10.1177/0363546506288239
62. Wang C, Ghalambor N, Zarins B, Warner JJP. Arthroscopic versus open Bankart repair: Analysis of patient subjective outcome and cost. *Arthroscopy*. 2005;21(10):1219–1222. doi:10.1016/j.arthro.2005.07.004
63. Kim SH, Ha KI, Kim SH. Bankart repair in traumatic anterior shoulder instability. *Arthroscopy*. 2002;18(7):755–763. doi:10.1053/jars.2002.31701
64. Sperber A, Hamberg P, Karlsson J, Swärd L, Wredmark T. Comparison of an arthroscopic and an open procedure for posttraumatic instability of the shoulder: A prospective, randomized multicenter study. *J Shoulder Elbow Surg*. 2001;10(2):105–108. doi:10.1067/mse.2001.112019
65. Karlsson J, Magnusson L, Ejerhed L, Hultenheim I, Lundin O, Kartus J. Comparison of open and arthroscopic stabilization for recurrent shoulder dislocation in patients with a Bankart lesion. *Am J Sports Med*. 2001;29(5):538–542. doi:10.1177/03635465010290050201
66. Geiger DF, Hurley JA, Tovey JA, Rao JP. Results of arthroscopic versus open Bankart suture repair. *Clin Orthop Relat Res*. 1997;337:111–117. doi:10.1097/00003086-199704000-00013
67. An VVG, Sivakumar BS, Phan K, Trantalis J. A systematic review and meta-analysis of clinical and patient-reported outcomes following two procedures for recurrent traumatic anterior instability of the shoulder: Latarjet procedure vs Bankart repair. *J Shoulder Elbow Surg*. 2016;25(5):853–863. doi:10.1016/j.jse.2015.11.001
68. Rollick N, Ono Y, Kurji HM, et al. Long-term outcomes of the Bankart and Latarjet repairs: A systematic review. *Open Access J Sports Med*. 2017;8:97–105. doi:10.2147/oajsm.s106983
69. Bliven KCH, Parr GP. Outcomes of the Latarjet procedure compared with Bankart repair for recurrent traumatic anterior shoulder instability. *J Athl Train*. 2018;53(2):181–183. doi:10.4085/1062-6050-232-16
70. Hurley ET, Jamal MS, Ali ZS, Montgomery C, Pauzenberger L, Mullett H. Long-term outcomes of the Latarjet procedure for anterior shoulder instability: A systematic review of studies at 10-year follow-up. *J Shoulder Elbow Surg*. 2019;28(2):e33–e39. doi:10.1016/j.jse.2018.08.028
71. Hurley ET, Lim Fat D, Farrington SK, Mullett H. Open versus arthroscopic Latarjet procedure for anterior shoulder instability: A systematic review and meta-analysis. *Am J Sports Med*. 2019;47(5):1248–1253. doi:10.1177/0363546518759540
72. Bøe B, Støen RØ, Blich I, Moatshe G, Ludvigsen TC. Learning curve for arthroscopic shoulder Latarjet procedure shows shorter operating time and fewer complications with experience. *Arthroscopy*. 2022;38(8):2391–2398. doi:10.1016/j.arthro.2022.01.042
73. Gendre P, Thélu CE, d'Ollonne T, Trojani C, Gonzalez JF, Boileau P. Coracoid bone block fixation with cortical buttons: An alternative to screw fixation? *Orthop Traumatol Surg Res*. 2016;102(8):983–987. doi:10.1016/j.otsr.2016.06.016
74. Valsamis EM, Kany J, Bonnevialle N, et al. The arthroscopic Latarjet: A multisurgeon learning curve analysis. *J Shoulder Elbow Surg*. 2020;29(4):681–688. doi:10.1016/j.jse.2019.10.022
75. Randelli P, Fossati C, Stoppani C, Evola FR, De Girolamo L. Open Latarjet versus arthroscopic Latarjet: Clinical results and cost analysis. *Knee Surg Sports Traumatol Arthrosc*. 2016;24(2):526–532. doi:10.1007/s00167-015-3978-9
76. Çağlar C, Akçaalan S, Akbulut B, Kengil MC, Uğurlu M, Doğan M. Open Latarjet reduces residual apprehension, re-dislocation and possibility of dislocation arthropathy compared to arthroscopic Bankart repair despite greater bipolar bone loss in anterior glenohumeral instability. *JSES Int*. 2024;8(6):1175–1181. doi:10.1016/j.jseint.2024.08.181
77. Vuletić F, Bøe B. Current trends and outcomes for open vs arthroscopic Latarjet. *Curr Rev Musculoskelet Med*. 2024;17(5):136–143. doi:10.1007/s12178-024-09889-9

Postpartum assessment of insulin resistance indicators in mothers with gestational diabetes: A prospective observational case-control study

Karolina Karcz[A–D,F], Paulina Gaweł[B,D], Barbara Królak-Olejnik[A,E,F]

Department of Neonatology, 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(5):779–794

Address for correspondence
Karolina Karcz
E-mail: k.karcz@umw.edu.pl

Funding sources
The language editing, proofreading and publication fees were funded by the Wroclaw Medical University. This work was supported as a project from the subsidy for maintaining and developing research potential in 2023. The research results presented here have been carried out within the framework of the theme, with funding from the subsidy No. SUBK.A300.23.024.

Conflict of interest
None declared

Acknowledgements
We are very grateful for the assistance of Barbara and Robyn Royle, native speakers of British English, in the English editing of the final version of the manuscript. Particular gratitude is given to Ms. Halina Biel-Milunovic, a retired clinical administrator at Keck School of Medicine, University of Southern California, for checking the accuracy of the English language and medical terminology. We would like to thank Ms. Lucyna Świderska for her help in collecting biological material.

Received on February 17, 2025
Reviewed on June 10, 2025
Accepted on July 15, 2025

Published online on March 16, 2026

Cite as
Karcz K, Gaweł P, Królak-Olejnik B. Postpartum assessment of insulin resistance indicators in mothers with gestational diabetes: A prospective observational case-control study. *Adv Clin Exp Med*. 2026;35(5):779–794. doi:10.17219/acem/208287

DOI
10.17219/acem/208287

Abstract

Background. Gestational diabetes mellitus (GDM) is characterized by a higher degree of insulin resistance (IR) than in a normal pregnancy. Several surrogate measures have been proposed to assess insulin sensitivity, including glycated hemoglobin, the Homeostatic Model Assessment–Insulin Resistance (HOMA-IR), and the Quantitative Insulin Sensitivity Check Index (QUICKI).

Objectives. The aim of the study was to determine whether markers of IR in the 1st week postpartum differ between mothers with GDM and healthy controls, and whether mothers with GDM treated with insulin have significantly different levels of IR markers compared with those treated with diet and physical activity. The association between IR markers, pregnancy outcomes, and maternal glucose profiles based on the oral glucose tolerance test (OGTT) was also investigated.

Materials and methods. Among the 70 participants, 50 mothers were diagnosed with GDM; 21 were treated with diet and physical activity (GDM G1), while 29 received insulin therapy (GDM G2). The remaining 20 participants constituted a control group with no history of glucose intolerance before or during pregnancy (non-GDM). A range of statistical methods (e.g., analysis of variance (ANOVA), Kruskal–Wallis test, χ^2 test, regression analysis, and cluster analysis) were used to compare data between study groups, with a significance level of $\alpha = 0.05$.

Results. The results showed that selected markers of IR in the 1st week after delivery differed significantly between mothers. Mothers with GDM exhibited considerably higher levels of HOMA-IR and HbA1c ($p < 0.05$), yet no substantially divergent QUICKI ($p > 0.05$) in the 1st week postpartum. Additionally, they demonstrated elevated glucose levels at 3 OGTT time points in comparison with non-GDM mothers. The GDM G2 group exhibited higher values than the GDM G1 group, except for the 1 h OGTT. Identification of maternal glucose phenotypes confirmed variability in the degree of glucose metabolism disorders among mothers.

Conclusions. Cluster analysis and glucose phenotype stratification in mothers with GDM help identify high-risk groups and support targeted counseling and monitoring to improve pregnancy outcomes.

Key words: insulin resistance, glycated hemoglobin, gestational diabetes, pregnancy outcomes

Highlights

- Gestational diabetes mellitus (GDM) is linked to maternal and neonatal complications.
- HOMA-IR, QUICKI, and glycated hemoglobin predict adverse pregnancy outcomes.
- Cluster and glucose phenotype analyses identify distinct GDM risk groups.
- Clinical use may improve risk stratification and targeted monitoring.
- Advanced GDM screening supports personalized care beyond binary diagnosis.

Background

Gestational diabetes mellitus (GDM) is a form of diabetes that manifests for the 1st time in a pregnant woman who was not previously diagnosed with diabetes. It can occur in any trimester of pregnancy. Nevertheless, its occurrence is more prevalent during the 2nd or 3rd trimester.[1,2] Gestational diabetes is not exclusively attributable to a paucity of insulin, but is also precipitated by insulin resistance (IR) or diminished insulin production by pancreatic beta cells. In pregnant women who demonstrate typical glucose metabolism, the compensatory mechanisms employed to address pregnancy-related insulin insensitivity and/or impaired β-cell function may be superseded by an adequate enhancement in insulin production. Whenever the above mechanism fails, GDM is likely to develop.[3–5] Despite the absence of a definitive understanding of the underlying causes of GDM, the prevailing hypothesis suggests that the presence of other hormones produced during pregnancy, or cytokines, can result in a reduction in the effectiveness of insulin, causing glucose to build up in the blood and be transported to the growing fetus instead of being absorbed by the maternal cells. Therefore, the presence of physiological IR during pregnancy has been shown to be advantageous for fetal growth and effective delivery of nutrients. However, if the degree of IR exceeds that which is typically observed in a normal pregnancy, then there are numerous adverse effects on the mother and fetus.[3,4] The symptoms of GDM usually disappear after delivery. Every year, GDM affects 5.4% of pregnancies in Europe and 3.4% in Poland.[5,6]

The Polish Society of Gynecologists and Obstetricians[6] adopted a set of diagnostic criteria for GDM, as outlined by the World Health Organization (WHO)[7] and the American Diabetes Association (ADA).[8] It is recommended that all pregnant women undergo measurement of fasting plasma glucose in the 1st trimester of pregnancy. The diagnosis is made on the basis of an oral glucose tolerance test (OGTT), with the administration of 75 g of glucose, and it is essential that only 1 of the following criteria is met: 1) fasting blood glucose of 92–125 mg/dL (5.6–6.9 mmol/L), 2) glycemia in 1-h OGTT ≥180 mg/dL (≥10.0 mmol/L), 3) blood glucose level in 2-h OGTT 153–199 mg/dL (8.5–11.1 mmol/L). Pregnant women without diagnosed pre-pregnancy diabetes should be screened for GDM at 24–28 weeks of gestation. If a pregnant woman is at higher risk of GDM (overweight or obesity, family history of diabetes, previous delivery of a macrosomic infant or an infant weighing more than 4,000 g, age over 25 years, race: African American, American Indian, Asian American, Hispanic or Latino, or Pacific Islander), screening should be done in the 1st trimester or immediately if her fasting glucose level is above the limits. It is recommended that women diagnosed with GDM be further evaluated for persistent diabetes between 6 and 12 weeks following childbirth.[6–8]

Special consideration should be given to overt diabetes first diagnosed in pregnancy. Overt diabetes meets the diagnostic criteria and the blood glucose cutoff values for diabetes in adults who are not pregnant. The International Association of Diabetes and Pregnancy Study Groups (IADPSG) has established a set of guidelines for the diagnosis of pre-pregnancy diabetes. These include the assessment of fasting plasma glucose concentration, 2-h plasma glucose during an OGTT with 75 g of glucose, and the percentage of hemoglobin A1c (HbA1c) ≥6.5% (48 mmol/mol). However, in the 2nd and 3rd trimesters of pregnancy, HbA1c may not be a useful test for the diagnosis of diabetes. Criteria for the diagnosis of diabetes in pregnancy are as follows: 1) 2 fasting plasma glucose values ≥126 mg/dL (7.0 mmol/L); 2) 2-h post-OGTT plasma glucose ≥200 mg/dL (11.1 mmol/L); 3) persistent glycemia ≥200 mg/dL (≥11.1 mmol/L) and associated hyperglycemic symptoms. Meeting one of these criteria is sufficient to establish a diagnosis. In addition, it is important to note that diagnostic criteria for diabetes based on blood glucose 1 h post-OGTT have not been established, and this value should not be used to diagnose diabetes in pregnancy.[2,7,8]

It is usually possible to manage and prevent the complications of GDM. However, the most common are hypertension in pregnancy, fetal macrosomia, increased risk of cesarean section, neonatal prematurity, respiratory distress, and hypoglycemia. Moreover, both the mother and the infant exhibit an augmented susceptibility to the development of type 2 diabetes in the future.[9–13] Therefore, the objective of GDM treatment is to maintain blood glucose concentrations within the normal range. The objective can be accomplished through a combination of dietary modifications, physical activity, daily self-assessment of blood glucose levels, and insulin administration.[6–8,10,12]

Objectives

The objective of this preliminary study was to investigate whether there are any differences in the levels of IR markers in the 1st 7 days after delivery between GDM mothers and healthy (non-GDM) controls. Furthermore, it was investigated whether GDM mothers receiving insulin treatment (GDM G2 group) exhibit significantly different levels of IR indicators compared to those managed with diet and physical activity (GDM G1 group). The association of IR markers with pregnancy outcomes was also investigated. In addition, it was also assessed whether the clusters obtained differ in basic parameters of neonatal body composition (birth weight, PI – ponderal index, TBW% – percentage of total body water, FBM – fat body mass, FBM% – percentage of fat body mass, and E/I ratio – extracellular/intracellular water ratio).

The secondary objective of the study was to identify subgroups of patients with GDM who were at a heightened risk of complications. This identification was based on a combination of both insulin-based and non-insulin-based markers and predictors.

The ultimate objective was to enhance individualized medical approaches by enabling early risk stratification and tailored interventions.

Materials and methods

Study design

The study was conceptualized as a prospective observational case-control study. Initiated in 2020, the study was suspended due to low enrolment during the COVID-19 pandemic. It was subsequently restarted in 2023. As demonstrated in the present article, the initial results of the indicators of IR were measured in a sample of 70 recruited mothers.

Inclusion and exclusion criteria

The following criteria were utilized to determine inclusion in the study groups: maternal age ranging from 18 to 45 years; delivery at or after 35 + 0/7 weeks of gestation by vaginal or caesarean route; singleton pregnancy; a newborn demonstrating a good postnatal condition, as evaluated by an Apgar score ≥7 points; and breastfeeding characterized by exclusivity or predominance. The recruitment process was initiated within the initial 7 postnatal days. Maternal consent for the assessment of IR markers was required. The exclusion criteria encompassed the following: any maternal or neonatal clinical condition that had an adverse effect on neonatal nutritional status, such as fetal growth retardation (FGR), absence of medical supervision during pregnancy, maternal use of alcohol or other psychoactive substances, nicotinism in pregnancy, uncontrolled maternal asthma, and the presence of congenital metabolic disease in the mother or neonate.

Study groups

In the present study, a total of 70 participants were included in the analysis, of whom 50 had a diagnosis of GDM. Of these, 21 were managed with a combination of diet and physical activity (GDM G1), while 29 received insulin therapy (GDM G2). The remaining 20 participants (non-GDM) were assigned to a control group. Prior to or during pregnancy, these individuals had no documented medical diagnosis of glucose intolerance. Additionally, their OGTT results in the current pregnancy were within normal limits. The enrollment took place among mothers hospitalized postpartum in the Second Department of Gynecology and Obstetrics, University Hospital in Wrocław (Poland), whose infants were hospitalized after birth in the Department of Neonatology of the University Hospital in Wrocław, Poland. Each mother was screened for GDM according to the recommendations of the Polish Society of Gynecologists and Obstetricians.[6]

Data collection

A retrospective analysis of maternal body weight during pregnancy was conducted using medical records. The categorization of gestational weight gain was determined in accordance with the guidelines published in the Committee Opinion of the American College of Obstetricians and Gynecologists, which were endorsed by the Polish Society of Gynecologists and Obstetricians.[6,14] The establishment of body mass index (BMI) ranges for underweight, normal weight, overweight, and obese women was based on maternal pre-pregnancy BMI.

The description of maternal weight gain in this study was as follows: below recommendations, within recommendations, or above recommendations. The collection of clinical data on the course of pregnancy, maternal antenatal history, labor, and puerperium was based on maternal medical records and personal interviews. The clinical data pertaining to the newborns were extracted from their medical records. The measurement of maternal and neonatal postpartum body weights was conducted using electronic medical scales, RADWAG type WPT 6/15D (RADWAG, Radom, Poland), to the nearest 10 g for neonates, and WPT 100/2000 (RADWAG), to the nearest 100 g for mothers.

Indicators of insulin resistance

To address the specific objectives of the project, insulin-based indicators of IR were employed. Peripheral venous blood for laboratory tests was taken from the mothers in the early morning after overnight fasting using the BD Vacutainer blood collecting system (Becton Dickinson, Franklin Lakes, USA). Fasting glucose [mg/dL], insulin [μU/mL] levels, and glycated hemoglobin (HbA1c; in %) were measured in the hospital

laboratory using Abbott analyzers (Abbott Laboratories, Chicago, USA). The following methods were used for the laboratory testing: the enzymatic method for the determination of glucose levels and immunoassays to measure insulin levels and HbA1c. The ensuing findings were employed to calculate 2 key indices: the Homeostatic Model Assessment–Insulin Resistance (HOMA-IR) and the Quantitative Insulin Sensitivity Check Index (QUICKI). The HOMA-IR was determined using the following formula: (fasting insulin [μIU/mL] × fasting glucose [mg/dL])/405. The QUICKI, on the other hand, was calculated as the inverse of the sum of the logarithms of fasting insulin and fasting glucose, expressed as: 1/(log(fasting insulin μIU/mL) + log(fasting glucose mg/dL)). The cutoff values for the diagnosis of IR were >2.0 for HOMA-IR and <0.34 for QUICKI. The target level of HbA1c was ≤6.1%.[15,16]

Glucose phenotypes

In their analysis, Yeung et al. examined both maternal and neonatal outcomes, with the maternal glucose phenotypes being identified from the results of a 50 g glucose challenge test (GCT) and a 75 g OGTT. They identified 7 glucose profiles that differed in the risk of perinatal complications.[17] The above method was used, differentiating mothers based on 75 g OGTT results only. The combinations of normal or elevated values of fasting glucose level (FGL), 1 h post-OGTT glucose (1-h POGL), and 2 h post-OGTT glucose (2-h POGL) were set, resulting in the identification of 8 phenotypes.

Neonatal body composition

A non-invasive method of bioimpedance analysis (BIA) was used to assess neonatal body composition, a process that involves the identification of distinct body compartments based on the electrical properties of human tissue. The analyses were conducted using the Body Composition Monitor (BCM; Fresenius Medical Care, Bad Homburg vor der Höhe, Germany) and specially designed disposable electrodes (BCM-FMC, <25 kg). The analysis was conducted at 50 frequencies ranging from 5 to 1,000 kHz, with an amplitude of 0.8 mA.[18] It should be noted that the measurements were performed by the same investigator, designated as K.K., in strict accordance with the manufacturer's instructions. The measurements were performed during the 1st week following birth, coinciding with the analysis of the mother's blood. For a comprehensive exposition of the methodology employed, the reader is referred to a separate manuscript which has been published independently.[19]

Microsoft Excel for Microsoft 365 (Microsoft, Redmond, USA), Statistica v. 13.3 (StatSoft, Inc., Tulsa, USA), and R v. 3.6.2 (R Foundation for Statistical Computing, Vienna, Austria)[20,21] were used for the statistical analysis.

Statistical analyses

The statistical analysis was conducted with a significance level set at $\alpha = 0.05$. A p-value less than 0.05 was considered statistically significant. The data were expressed as the mean and standard deviation (SD), median and interquartile range (IQR), or number of cases (n) and percentage, as appropriate. One-way analysis of variance (ANOVA), the Kruskal–Wallis test, or the χ^2 test was used to compare the data between the study groups, depending on the type of data and its distribution. The Shapiro–Wilk test was employed to ascertain the normality of the data set. Levene's test was conducted to evaluate the homogeneity of variances across variables. Supplementary Table 1 presents the results of the aforementioned tests.

The impact of specific maternal characteristics on postpartum indicators of IR was evaluated with the use of univariate regression (generalized linear model). Subsequently, cluster analysis was conducted using the Marczewski–Steinhaus (M–S) taxonomic approach, and a dendrogram was constructed.[22] The impact of the resulting clusters was ultimately determined through the implementation of 2 analytical methods: a one-way ANOVA and the Kruskal–Wallis test. To ensure the reliability of the established taxonomic approach, the expectation–maximization (E–M) algorithm was employed for validation.[23]

A dendrogram is a chart that graphically represents the results of a cluster analysis. Clusters are groups of patients or data with maximal within-group similarities and maximal between-group differences in the investigated variables. The construction of a dendrogram facilitates the visualization of the formation of clusters at each stage and the assessment of the degree of similarity (or distance) among the clusters. According to Marczewski and Steinhaus, the distance is equal to the differences between patients in the variables studied. Once the distance between variables is defined, linkage methods are used to link sufficiently similar clusters.[22,23] The E–M algorithm is a maximum likelihood estimation method for latent variable models that is considered to have a high degree of reliability and complexity.[23]

Ethics

The study was conducted in compliance with the Helsinki Declaration and received approval from the Bioethics Committee of the Medical University of Wroclaw, Poland. The protocol (code KB 950/2022) was authorized on December 21, 2022, as a continuation of KB 407/2020 from June 23, 2020, KB 35/2020 from January 16, 2020, and 773/2019 from November 25, 2019. The mothers who participated in the study groups were given the opportunity to provide their written informed consent. The current study has been duly registered in the ClinicalTrials.gov registry and was assigned the number NCT04937348.

Results

General description of the study group

The mean maternal age was 32.7 (SD = 4.5) years. The median gravidity was 2.0 (IQR = 1.0), and parity was also 2.0 (IQR = 1.0). A review of the obstetric records of the participating mothers resulted in the following findings: 6 (8.6%) had been diagnosed with GDM in a previous pregnancy, 7 (10%) delivered a macrosomic infant (>90th percentile for birth weight for gestational age), and 6 (8.6%) had newborns weighing more than 4 kg. These characteristics were not statistically different between the study groups ($p > 0.05$). Pregestational diabetes or IR had not been previously diagnosed in any of these mothers.

Based on the medical history and outcomes of previous pregnancies, the following results were obtained. Most of the newborns were girls (60%), and the most common method of delivery was cesarean section (over 70%). The median number of weeks of gestation at birth was 39 (IQR = 2.0). Mean neonatal anthropometric measurements at birth were as follows: birth weight 3.5 (SD = 0.5) kg, length 53.2 (SD = 2.8) cm, and head circumference 34.7 (SD = 1.5) cm. The median percentile for birth weight was 66.0 (IQR = 44.0).

Considering classification by neonatal weight for gestational age, 80% (n = 56) of newborns were eutrophic (AGA – appropriate for gestational age), 4.3% (n = 3) were hypotrophic (SGA – small for gestational age), and 15.7% (n = 11) were macrosomic (LGA – large for gestational age). Six newborns (8.6%) developed respiratory distress after birth, including 4 born to GDM G2 mothers. These results were not statistically different between the analyzed groups ($p > 0.05$). Significant differences were observed in the incidence of congenital malformations in newborns; all 4 cases occurred in the GDM G1 group ($\chi^2(2, n = 70) = 9.899$, $p = 0.006$; $V = 0.376$). In some of these newborns, more than one defect was identified. Among these newborns, 3 had hydronephrosis, 1 had a ventricular septal defect (VSD), 1 had an atrial septal defect type II (ASD II), 1 had a persistent left superior vena cava (LVCS), and 1 had a cleft soft palate.

Newborns of GDM mothers were mostly fed predominantly with breast milk and occasionally supplemented with formula (13 (61.9%) in the GDM G1 group and 19 (65.5%) in the GDM G2 group), whereas newborns of non-GDM mothers were mainly exclusively breastfed (15 (75%)). These results were statistically significant (χ^2 (2, n = 70) = 8.784, $p = 0.012$; $V = 0.354$). Maternal morbidity was similar between the studied groups ($p > 0.05$), including the incidence of hypertension, hypothyroidism, nicotine use before pregnancy, nausea and vomiting in the 1st trimester, and rates of vaginal or urinary tract infections.

The detailed results of the aforementioned characteristics with comparisons between the study groups are shown in Table 1.

Maternal nutritional status, glucose screening tests, and markers of insulin resistance

The mothers differed significantly between the study groups in terms of pre-gestational BMI ($H(2, n = 70) = 8.537$, $p = 0.014$) and classification of nutritional status according

Table 1. Comparison of selected characteristics and pregnancy outcomes across study groups

Parameter		All participants (n = 70)	GDM G1 (n = 21)	GDM G2 (n = 29)	non-GDM (n = 20)	p-value
Selected maternal characteristics						
Maternal age [years], mean (SD)		32.7 (4.5)	33.9 (4.9)	32.1 (4.6)	32.5 (3.9)	0.363[b]
Maternal pre-gestational BMI, median (IQR)		24.2 (6.4)	23.3 (3.5)	28.0 (6.8)	23.0 (2.8)	0.014[a]
Maternal classification by pre-gestational BMI, n (%)	normal	40 (57.1)	14 (66.7)	9 (31.0)	17 (85.0)	0.004[c]
	overweight	16 (22.9)	4 (19.0)	11 (38.0)	1 (5.0)	
	obese	14 (20.0)	3 (14.3)	9 (31.0)	2 (10.0)	
Maternal weight gain during pregnancy [kg], mean (SD)		11.5 (5.7)	10.2 (3.5)	9.2 (5.4)	16.2 (5.4)	<0.001[b]
Maternal weight gain during pregnancy in reference to pre-gestational BMI, n (%)	below recommendations	18 (25.8)	7 (33.3)	9 (31.0)	2 (10.0)	0.033[c]
	within recommendations	26 (37.1)	10 (47.7)	11 (38.0)	5 (25.0)	
	above recommendations	26 (37.1)	4 (19.0)	9 (31.0)	13 (65.0)	
Fasting glucose in the 1st trimester of pregnancy [mg/dL], median (IQR)		89.0 (14.2)	89.0 (17.0)	93.0 (9.2)	85.6 (18.9)	0.123[a]
Fasting glucose in the 2nd/3rd trimester of pregnancy [mg/dL], median (IQR)		92.0 (15.0)	90.0 (15.0)	95.0 (10.0)	85.3 (18.0)	0.003[a]
1-h post OGTT glucose level [mg/dL], median (IQR)		135.5 (82.0)	125.0 (81.0)	182.0 (46.0)	108.5 (32.7)	<0.001[a]
2-h post OGTT glucose level [mg/dL], median (IQR)		119.5 (50.0)	138.0 (53.0)	130.0 (47.0)	102.0 (12.0)	0.001[a]
Maternal HbA1c [%] in the 1st week postpartum, median (IQR)		5.4 (0.4)	5.3 (0.4)	5.6 (0.3)	5.3 (0.5)	0.009[a]
Maternal HOMA-IR in the 1st week postpartum, median (IQR)		0.66 (1.67)	0.55 (0.35)	0.80 (0.83)	0.49 (0.55)	0.046[a]
Maternal QUICKI in the 1st week postpartum, median (IQR)		0.41 (0.13)	0.43 (0.05)	0.41 (0.06)	0.43 (0.07)	0.106[a]

Table 1. Comparison of selected characteristics and pregnancy outcomes across study groups – cont.

Parameter		All participants (n = 70)	GDM G1 (n = 21)	GDM G2 (n = 29)	non-GDM (n = 20)	p-value
Characteristics referring to a newborn infant						
Pregnancy duration [weeks], median (IQR)		39.0 (2.0)	39.0 (2.0)	38.0 (1.0)	39.0 (2.0)	0.197[a]
Gender of newborn, n (%)	boys	28 (40.0)	7 (33.3)	12 (41.4)	9 (45.0)	0.733[c]
	girls	52 (60.0)	14 (66.7)	17 (58.6)	11 (55.0)	
Mode of delivery, n (%)	vaginal birth	20 (28.6)	7 (33.3)	7 (24.1)	6 (30.0)	0.766[c]
	cesarean section	50 (71.4)	14 (66.7)	22 (75.9)	14 (70.0)	
Newborn's birth weight [kg], mean (SD)		3.5 (0.5)	3.4 (0.6)	3.3 (0.4)	3.7 (0.4)	0.079[b]
Newborn's length [cm], mean (SD)		53.2 (2.8)	53.0 (3.5)	53.1 (2.5)	53.6 (2.3)	0.767[b]
Newborn's head circumference [cm], mean (SD)		34.7 (1.5)	34.5 (1.6)	34.8 (1.6)	34.7 (1.6)	0.887[b]
Percentile for neonatal birth weight, median (IQR)		66.0 (44.0)	56.0 (45.0)	69.0 (49.0)	75.0 (35.5)	0.363[a]
Classification of neonatal birth weight, n (%)	SGA	3 (4.3)	1 (4.7)	2 (6.9)	0 (0.0)	0.657[c]
	AGA	56 (80.0)	17 (81.0)	23 (79.3)	16 (80.0)	
	LGA	11 (15.7)	4 (14.3)	3 (13.8)	4 (20.0)	
Neonatal respiratory distress after birth, n (%)		6 (8.6)	1 (4.8)	4 (13.8)	1 (5.0)	0.422[c]
Congenital malformation in neonate, n (%)		4 (5.7)	4 (19.1)	0 (0.0)	0 (0.0)	0.006[c]
Method of feeding of a newborn, n (%)	exclusive breastfeeding	33 (47.1)	8 (38.1)	10 (34.5)	15 (75.0)	0.012[c]
	mixed feeding	37 (52.9)	13 (61.9)	19 (65.5)	5 (25.0)	
Maternal medical history during and before pregnancy						
Maternal hypertension, n (%)	chronic (onset before the pregnancy)	6 (8.6)	1 (4.8)	5 (17.2)	0 (0.0)	0.478[c]
	pregnancy-induced	8 (11.4)	2 (9.5)	5 (17.2)	1 (5.0)	
	none	56 (80.0)	18 (85.7)	19 (65.6)	19 (95.0)	
Maternal hypothyroidism, n (%)	chronic (onset before the pregnancy)	20 (28.6)	6 (28.6)	12 (41.4)	2 (10.0)	0.175[c]
	gestational (onset during the pregnancy)	13 (18.6)	5 (23.8)	4 (13.8)	4 (20.0)	
	none	37 (52.9)	10 (47.6)	13 (44.8)	14 (70.0)	
Nicotinism before pregnancy, n (%)		22 (31.4)	6 (28.6)	11 (37.9)	5 (25.0)	0.597[c]
Nausea and vomiting, n (%)		14 (20.0)	5 (23.8)	6 (20.7)	3 (15.0)	0.774[c]
Initially excessive weight loss, n (%)		14 (20.0)	6 (28.6)	7 (24.0)	1 (5.0)	0.085[c]
VI, n (%)		6 (8.6)	1 (3.5)	4 (19.1)	1 (5.0)	0.121[c]
UTI, n (%)		6 (8.6)	3 (14.3)	3 (10.3)	0 (0.0)	0.239[c]
Obstetric history						
Gravidity, median (IQR)		2.0 (1.0)	2.0 (2.0)	2.0 (1.0)	2.0 (1.0)	0.520[a]
Parity, median (IQR)		2.0 (1.0)	1.0 (1.0)	1.0 (1.0)	2.0 (1.0)	0.395[a]
Developed GDM in previous pregnancy, n (%)		6 (8.6)	3 (14.3)	3 (10.3)	0 (0.0)	0.238[c]
Delivered a neonate with body weight >4 kg in previous pregnancy, n (%)		6 (8.6)	3 (14.3)	1 (3.5)	2 (10.0)	0.387[c]
Delivered a hypertrophic neonate in previous pregnancy, n (%)		7 (10.0)	2 (9.5)	2 (6.9)	3 (15.0)	0.657[c]
Delivered a hypotrophic neonate in previous pregnancy, n (%)		0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1.0[c]

[a] – Kruskal–Wallis test; [b] – one-way analysis of variance (ANOVA); [c] – χ^2 test; SD – standard deviation; IQR – interquartile range; n – number of patients; GDM – gestational diabetes mellitus; non-GDM – control group; M G1 – group of mothers with GDM who were administered diet and physical activity; GDM G2 – group of mothers with GDM receiving insulin treatment; BMI – body mass index; OGTT – oral glucose tolerance test; HbA1c – hemoglobin A1c; HOMA-IR – Homeostatic Model Assessment–Insulin Resistance; QUICKI – Quantitative Insulin Sensitivity Check Index; SGA – small for gestational age; AGA – appropriate for gestational age; LGA – large for gestational age; VI – vaginal infections; UTI – urinary tract infections.

to pre-gestational BMI ($\chi^2(4, n = 70) = 15.424$, $p = 0.004$; $V = 0.332$). The highest pre-gestational BMI was found in the GDM G2 group, with a median of 28.0 (IQR = 6.8) kg/m^2. In this group, the highest percentage of mothers were overweight or obese compared with the other study groups. However, the highest weight gain during pregnancy was observed in non-GDM mothers, with a mean of 16.2 (SD = 5.4) kg ($F(2, 67) = 12.923$, $p < 0.001$ for the study groups), which corresponded to the highest incidence (13 (65%)) of gestational weight gain above obstetrical recommendations ($\chi^2(4, n = 70) = 10.454$, $p = 0.033$; $V = 0.273$).

In the 1st trimester of pregnancy, the median fasting glucose level was 89.0 mg/dL (IQR = 14.2) across all study groups. Subsequent analysis revealed no statistically

significant differences between the study groups with respect to this parameter. However, statistically significant differences ($p < 0.05$) were identified among the study groups with regard to OGTT results. The highest fasting glucose level observed during the 2nd or 3rd trimester (median 95.0 (IQR = 10.0) mg/dL) and the 1 h post-OGTT glucose level (median 182.0 (IQR = 46.0) mg/dL) were found in the GDM G2 group, whereas the highest median 2 h post-OGTT glucose level (median 138.0 (IQR = 53.0) mg/dL) was found in the GDM G1 group.

Regarding abnormal glucose screening results, elevated levels of the following parameters were observed: 1st-trimester fasting glucose (92–125 mg/dL) in 19 (27.1%) mothers, 2nd/3rd-trimester fasting glucose (92–125 mg/dL) in 30 (42.9%), 1 h post-OGTT glucose (≥180 mg/dL) in 25 (35.1%), and 2 h post-OGTT glucose (153–199 mg/dL) in 25 (35.1%) (Table 1). A comparison of the results obtained from the study groups revealed a significant difference in the levels of IR markers in the week following childbirth, including HbA1c ($H(2, n = 70) = 9.372$, $p = 0.009$) and HOMA-IR ($H(2, n = 70) = 6.177$, $p = 0.046$), with the highest values observed in the GDM G2 group. QUICKI values were similar ($p > 0.05$) between the study groups, with a median of 0.41 (IQR = 0.13) for all participants. An elevated HbA1c level (≥6.1%) was found in 2 (2.9%) participants. Elevated HOMA-IR (>2.0) and decreased QUICKI (<0.34) were found in 9 (12.9%) and 9 (12.9%) participants, respectively (Table 1).

There were 6 GDM mothers who had been diagnosed with GDM in a previous pregnancy. Compared with mothers diagnosed with GDM for the 1st time, these mothers had similar levels of HbA1c ($H(1, n = 21) = 0.510$, $p = 0.475$ for GDM G1 and $H(1, n = 29) = 0.033$, $p = 0.856$ for GDM G2), HOMA-IR ($H(1, n = 21) = 0.010$, $p = 0.920$ for GDM G1 and $H(1, n = 29) = 0.621$, $p = 0.432$ for GDM G2), and QUICKI ($H(1, n = 21) = 0.064$, $p = 0.800$ for GDM G1 and $H(1, n = 29) = 0.743$, $p = 0.389$ for GDM G2). There was no history of GDM in previous pregnancies among mothers in the non-GDM group.

Maternal factors and results of IR markers

The role of maternal characteristics in determining levels of IR indicators during the initial week postpartum was assessed using univariate regression, specifically a generalized linear model. In the analysis, the following factors were taken into account: study group (indicative of the extent of glucose intolerance), maternal chronological age, parity (number of times a woman has given birth), gravidity (total number of pregnancies regardless of duration and outcome), pre-pregnancy BMI, nutritional status based on pre-pregnancy BMI classification, maximum BMI in pregnancy, increase in body weight during pregnancy, classification of weight gain based on obstetric recommendations, body weight and BMI at recruitment after delivery, and medical interview data on hypertension, hypothyroidism, nicotinism, and results of glucose sensitivity screening tests conducted during pregnancy. The following factors were identified as the strongest determinants of IR markers during the initial postpartum period, as determined by the Akaike Information Criterion (AIC) values: allocation to a specific study group, maternal BMI before pregnancy, and maternal weight gain during pregnancy. A comprehensive overview of the outcomes is provided in Supplementary Table 2.

Cluster analysis

A classification tree for the patients was constructed based on the identified factors, including allocation to a specific study group, maternal pre-pregnancy BMI, and maternal gestational weight gain (Fig. 1).

The dendrogram shows 3 distinct patient types, the characteristics of which are presented in Table 2. The average M–S distance between pairs of compared sets (clusters vs original study groups) was approx. 0.027.

'Cluster 1' included GDM-affected mothers (n = 29 GDM G2 and n = 1 GDM G1), with the highest mean preconceptional BMI (mean 28.8 kg/m^2, SD = 0.9, corresponding

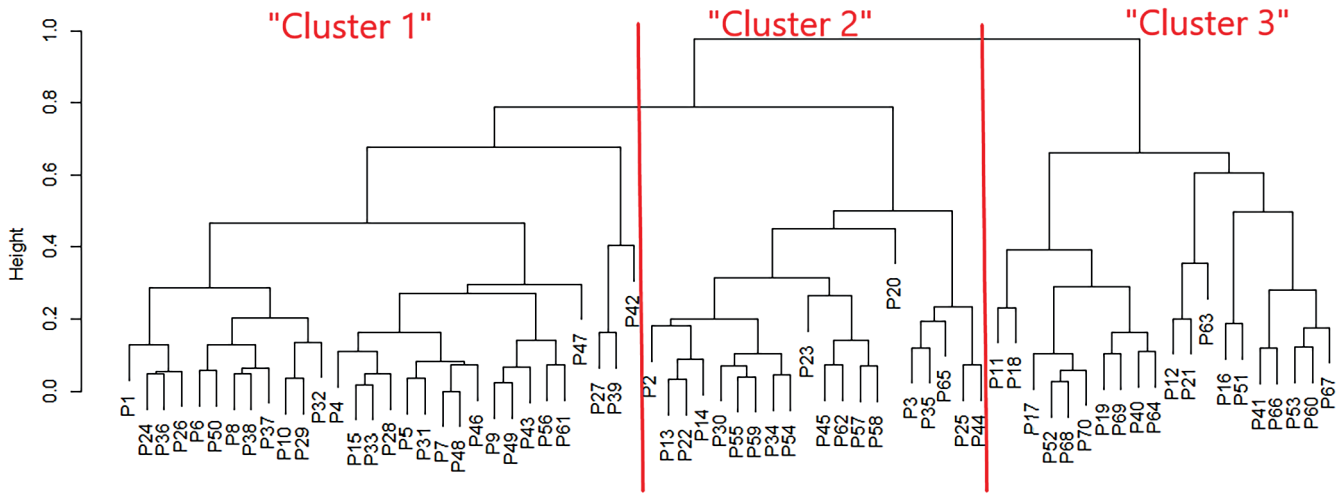

Fig. 1. Dendrogram showing the clustering of mothers

Table 2. Comparison of selected characteristics and pregnancy outcomes across clusters

Parameter		Cluster 1 (n = 30)	Cluster 2 (n = 20)	Cluster 3 (n = 20)	p-value
Selected maternal characteristics					
Maternal age [years], mean (SD)		32.2 (4.6)	33.9 (5.1)	32.5 (3.9)	0.429[b]
Maternal pre-gestational BMI, median (IQR)		28.1 (9.6)	23.2 (3.1)	22.96 (2.83)	0.002[a]
Maternal classification by pre-gestational BMI, n (%)	normal	9 (30.0)	14 (70.0)	17 (85.0)	0.002[c]
	overweight	11 (36.7)	4 (20.0)	1 (5.0)	
	obese	10 (33.3)	2 (10.0)	2 (10.0)	
Maternal weight gain during pregnancy [kg], mean (SD)		9.0 (5.4)	10.6 (3.2)	16.2 (5.4)	<0.001[b]
Maternal weight gain during pregnancy in reference to pre-gestational BMI, n (%)	below recommendations	10 (33.3)	6 (30.0)	2 (10.0)	0.031[c]
	within recommendations	11 (36.7)	10 (50.0)	5 (25.0)	
	above recommendations	9 (30.0)	4 (20.0)	13 (65.0)	
Fasting glucose in the 1st trimester of pregnancy [mg/dL], median (IQR)		92.5 (10.0)	99.0 (17.1)	85.6 (18.9)	0.132[a]
Fasting glucose in the 2nd/3rd trimester of pregnancy [mg/dL], median (IQR)		94.8 (11.0)	90.5 (14.0)	85.3 (18.0)	0.007[a]
1-hour post OGTT glucose level [mg/dL], median (IQR)		182.9 (46.0)	124.5 (81.0)	108.5 (32.7)	<0.001[a]
2-hour post OGTT glucose level [mg/dL], median (IQR)		131.5 (47.0)	134.5 (53.5)	102.0 (12.0)	0.001[a]
Maternal HbA1c [%] in the 1st week postpartum, median (IQR)		5.6 (0.3)	5.3 (0.5)	5.3 (0.5)	0.004[a]
Maternal HOMA-IR in the 1st week postpartum, median (IQR)		0.83 (0.83)	0.54 (0.33)	0.59 (0.55)	0.031[a]
Maternal QUICKI in the 1st week postpartum, median (IQR)		0.41 (0.06)	0.33 (0.05)	0.43 (0.07)	0.070[a]
Characteristics referring to a newborn infant					
Pregnancy duration [weeks], median (IQR)		38.6 (0.9)	38.8 (1.0)	39.0 (2.0)	0.293[a]
Gender of newborn, n (%)	boys	13 (43.3)	6 (30.0)	9 (45.0)	0.554[c]
	girls	17 (56.7)	14 (70.0)	11 (55.0)	
Mode of delivery, n (%)	vaginal birth	8 (26.7)	6 (30.0)	6 (30.0)	0.954[c]
	cesarean section	22 (73.3)	14 (70.0)	14 (70.0)	
Newborn's birth weight [kg], mean (SD)		3.3 (0.4)	3.4 (0.6)	3.65 (0.4)	0.077[b]
Newborn's length [cm], mean (SD)		53.1 (2.5)	53.0 (4.0)	53.6 (2.3)	0.767[b]
Percentile for neonatal birth weight, median (IQR)		57.0 (29.0)	57.0 (46.0)	74.5 (35.5)	0.349[a]
Classification of neonatal birth weight, n (%)	SGA	2 (6.7)	1 (5.0)	0 (0.0)	0.809[c]
	AGA	24 (80.0)	16 (80.0)	16 (80.0)	
	LGA	4 (13.3)	3 (15.0)	4 (20.0)	
Neonatal respiratory distress after birth, n (%)		4 (13.3)	1 (5.0)	1 (5.0)	0.468[c]
Congenital malformation in neonate, n (%)		1 (3.3)	3 (15.0)	0 (0.0)	0.094[c]
Method of feeding of a newborn, n (%)	exclusive breastfeeding	10 (33.3)	8 (40.0)	15 (75.0)	0.011[c]
	mixed feeding	20 (66.7)	12 (60.0)	5 (25.0)	
Maternal medical history during and before pregnancy					
Maternal hypertension, n (%)	chronic (onset before the pregnancy)	2 (6.6)	1 (5.0)	0 (0.0)	0.519[c]
	pregnancy-induced	5 (16.7)	2 (10.0)	1 (5.0)	
	none	23 (76.7)	17 (85.0)	19 (95.0)	
Maternal hypothyroidism, n (%)	chronic (onset before the pregnancy)	12 (28.6)	6 (41.4)	2 (10.0)	0.234[c]
	gestational (onset during the pregnancy)	5 (23.8)	4 (13.8)	4 (20.0)	
	none	13 (47.6)	10 (44.8)	14 (70.0)	
Nicotinism before pregnancy, n (%)		11 (36.7)	6 (30.0)	5 (25.0)	0.676[c]
Nausea and vomiting, n (%)		6 (20.0)	5 (25.0)	3 (15.0)	0.732[c]
Initially excessive weight loss, n (%)		8 (26.7)	5 (25.0)	1 (5.0)	0.089[c]
VI, n (%)		1 (3.3)	4 (20.0)	1 (5.0)	0.095[c]
UTI, n (%)		3 (10.0)	3 (15.0)	0 (0.0)	0.222[c]

Table 2. Comparison of selected characteristics and pregnancy outcomes across clusters – cont.

Parameter	Cluster 1 (n = 30)	Cluster 2 (n = 20)	Cluster 3 (n = 20)	p-value
Obstetric history				
Gravidity, median (IQR)	1.8 (0.9)	2.2 (1.3)	2.0 (1.0)	0.709[a]
Parity, median (IQR)	1.6 (0.9)	1.7 (1.0)	2.0 (1.0)	0.397[a]
Developed GDM in previous pregnancy, n (%)	3 (10.0)	3 (15.0)	0 (0.0)	0.222[c]
Delivered a neonate with body weight >4 kg in previous pregnancy, n (%)	3 (14.3)	1 (3.5)	2 (10.0)	0.387[c]
Delivered a hypertrophic neonate in previous pregnancy, n (%)	2 (9.5)	2 (6.9)	3 (15.0)	0.657[c]
Delivered a hypotrophic neonate in previous pregnancy, n (%)	0 (0.0)	0 (0.0)	0 (0.0)	1.0[c]

[a] – Kruskal–Wallis test; [b] – one-way analysis of variance (ANOVA); [c] – χ^2 test; SD – standard deviation; IQR – interquartile range; n – number of patients; GDM – gestational diabetes mellitus; non-GDM – control group; M G1 – group of mothers with GDM who were administered diet and physical activity; GDM G2 – group of mothers with GDM receiving insulin treatment; BMI – body mass index; OGTT – oral glucose tolerance test; HbA1c – hemoglobin A1c; HOMA-IR – Homeostatic Model Assessment–Insulin Resistance; QUICKI – Quantitative Insulin Sensitivity Check Index; SGA – small for gestational age; AGA – appropriate for gestational age; LGA – large for gestational age; VI – vaginal infections; UTI – urinary tract infections.

Table 3. Characteristics of mothers across clusters based on one-way analysis of variance (ANOVA)

Cluster	n	Study group			Maternal BMI before conception [kg/m^2], mean ±SD	Weight gain during pregnancy [kg], mean ±SD
		GDM G1, n	GDM G2, n	non-GDM, n		
1	30	1	29	0	28.8 ±0.9	9.0 ±0.9
2	20	20	0	0	24.4 ±1.1	10.6 ±1.1
3	20	0	0	20	24.1 ±1.1	16.2 ±1.1
F statistic	–	N/A			7.06	13.41
p-value	–	N/A			0.002	<0.001

N/A – non applicable; BMI – body mass index; SD – standard deviation; GDM – gestational diabetes mellitus. The F statistic of analysis of variance (ANOVA) was used to compare BMI clusters as measured before pregnancy.

Table 4. Multiple comparisons between clusters (post hoc Tukey–Kramer and Dunn's test p-values) for selected maternal insulin resistance (IR) markers and neonatal anthropometric and body composition parameters

Difference between clusters	Maternal HbA1c [%]	Maternal HOMA-IR	Maternal QUICKI	Neonatal birth weight [kg]	Neonatal Ponderal index [kg/m^3]	Neonatal TBW%	E/I	Neonatal FBM [kg]	Neonatal FBM%
1–2	0.017	0.093	0.19	1.0	0.579	0.676	0.8	0.397	0.353
1–3	0.015	0.073	0.14	0.096	0.123	0.316	1.0	0.069	0.136
2–3	1.0	1.0	1.0	0.334	0.642	0.841	1.0	1.0	0.870

HbA1c – hemoglobin A1c; HOMA-IR – Homeostatic Model Assessment–Insulin Resistance; QUICKI – Quantitative Insulin Sensitivity Check Index; TBW – total body water; E/I – extracellular/intracellular water ratio; FBM – fat body mass.

to a median of 28.1 kg/m^2, IQR = 9.6) and the lowest gestational weight gain (mean 5.4 kg, SD = 9.0) 'Cluster 2' included the remaining GDM G1 mothers, with a mean BMI before pregnancy of 24.1 kg/m^2 (SD = 1.1) corresponding to median 23.2 kg/m^2 (IQR = 3.1) and a mean weight gain during pregnancy of 10.6 kg (SD = 1.1). 'Cluster 3' comprised healthy mothers (non-GDM), with a mean BMI before pregnancy of 24.1 kg/m^2 (SD = 1.1), corresponding to median 22.96 kg/m^2 (IQR = 2.83) and the highest gestational weight gain (mean 16.2 kg, SD = 1.1) (Table 2).

As the cluster classification closely corresponded to the classification of mothers by study group, the results of the comparison of selected characteristics were similar. Mothers in the clusters differed significantly according to the classification of nutritional status based on pre-gestational BMI (χ^2 (4, n = 70) = 17.194, p = 0.002; V = 0.350), gestational weight gain according to obstetrical recommendations (χ^2(4, n = 70) = 10.599, p = 0.031; V = 0.275), and the method of infant feeding (χ^2 (2, n = 70) = 8.934, p = 0.011; V = 0.357). The statistical significance of the comparison of characteristics across the research groups was identical to that observed in the cluster comparison. The results of the comparison between clusters are summarized in Table 3.

When maternal IR markers were compared between clusters, a significant difference was observed in the Kruskal–Wallis test for HbA1c and HOMA-IR (H(2, n = 70) = 11.222, p = 0.004 and H(2, n = 70) = 6.950, p = 0.031, respectively). With respect to neonatal nutritional status and body composition, the clusters did not differ (p > 0.05). In multiple comparisons between clusters, no differences were observed in newborns' body composition or anthropometric assessment (p > 0.05). When maternal IR markers in the 1st week

postpartum were considered, 'Cluster 1' differed from 'Cluster 2' and from 'Cluster 3' in terms of HbA1c ($p < 0.05$) (Table 4).

Neonatal body composition in clusters

Clustered analysis did not reveal any significant differences regarding neonatal birth weight and body composition. Results of one-way ANOVA were as follows: $F(2, 67) = 2.666$, $p = 0.077$ for birth weight, $F(2, 67) = 1.963$, $p = 0.148$ for ponderal index, $F(2, 67) = 3.014$, $p = 0.056$ for FBM, $F(2, 67) = 2.099$, $p = 0.131$ for FBM%, and $F(2, 67) = 1.11$, $p = 0.336$ for TBW%. Kruskal–Wallis ANOVA was $H(2, n = 70) = 1.285$, $p = 0.526$ for E/I. Results in 'Cluster 1' were: mean birth weight 3.42 kg (SD 0.59), mean ponderal index 2.16 kg/m^2 (SD = 0.27), mean FBM 0.27 kg (SD = 0.09), mean FBM% 0.08 (SD = 0.02), mean TBW% 0.81 (SD 0.05), and median E/I 0.50 (IQR = 0.12). Results in 'Cluster 2' were: mean birth weight 3.34 kg (SD = 0.43), mean ponderal index 2.09 kg/m^2 (SD = 0.25), mean FBM 0.24 kg (SD = 0.09), mean FBM% 0.08 (SD = 0.02), mean TBW% 0.82 (SD = 0.07), and median E/I 0.44 (IQR = 0.08). Results in 'Cluster 3' were: mean birth weight 3.65 kg (SD = 0.40), mean ponderal index 2.24 kg/m^2 (SD = 0.28), mean FBM 0.31 kg (SD = 0.10), mean FBM% 0.09 (SD = 0.02), mean TBW% 0.80 (SD = 0.07), and median E/I 0.50 (IQR = 0.09).

Glucose phenotypes

Eight maternal glucose phenotypes were identified by analyzing OGTT at 3 time points: 1) normal OGTT; 2) elevated FGL + normal 1-h POGL + normal 2-h POGL; 3) normal FGL + elevated 1-h POGL + normal 2-h POGL; 4) normal FGL + normal 1-h POGL + elevated 2-h POGL; 5) elevated FGL + elevated 1-h POGL + elevated 2-h POGL; 6) elevated FGL + elevated 1-h POGL + normal 2-h POGL; 7) elevated FGL + normal 1-h POGL + elevated 2-h POGL; and 8) normal FGL + elevated 1-h POGL + elevated 2-h POGL. Normal OGTT profile included only non-GDM mothers, whereas both GDM G1 and GDM G2 mothers were present in the other 7 profiles in varying proportions ($p < 0.05$). The glucose phenotypes differed significantly in terms of maternal classification by pregestational BMI ($p = 0.014$) and gestational weight gain in relation to BMI before conception ($p = 0.018$). Mothers who had abnormal glucose levels at 3 time points of OGTT were found with the highest incidence of BMI > 25 kg/m^2, including the highest percentage of obesity and excessive gestational weight gain among assigned mothers. These mothers were also found to have the highest incidence of elevated HbA1c (≥6.1%) in the 1st week postpartum, hypothyroidism, urinary tract infections (UTI) during pregnancy, nicotine use before the pregnancy, as well as the highest incidence of delivering LGA newborns. However, these results were not statistically significant ($p > 0.05$). The other significant differences ($p < 0.05$) included the incidence of respiratory distress after birth in newborns (with the highest percentage of cases in elevated FGL + normal 1-h POGL + elevated 2-h POGL profile) and method of feeding a neonate (with the highest percentage of sporadic formula use in normal FGL + elevated 1-h POGL + normal 2-h POGL and elevated FGL + elevated 1-h POGL + normal 2-h POGL profiles). The detailed results are presented in Table 5.

Table 5. Comparison of glucose phenotypes across study groups

Selected gestational and perinatal outcomes	Glucose phenotypes based on OGTT								p-value
	normal OGTT (1)	elevated FGL + normal 1-h POGL + normal 2-h POGL (2)	normal FGL + elevated 1-h POGL + normal 2-h POGL (3)	normal FGL + normal 1-h POGL + elevated 2-h POGL (4)	elevated FGL + elevated 1-h POGL + elevated 2-h POGL (5)	elevated FGL + elevated 1-h POGL + normal 2-h POGL (6)	elevated FGL + normal 1h POGL + elevated 2-h POGL (7)	normal FGL + elevated 1h POGL + elevated 2-h POGL (8)	
	n = 20	n = 8	n = 6	n = 7	n = 8	n = 11	n = 3	n = 7	
Study group, n (%)									
GDM G1	0 (0.0)	3 (37.5)	3 (50.0)	4 (57.1)	2 (25.0)	2 (18.2)	2 (66.7)	5 (71.4)	<0.001[c]
GDM G2	0 (0.0)	5 (62.5)	3 (50.0)	3 (42.9)	6 (75.0)	9 (81.8)	1 (33.3)	2 (28.6)	
Non-GDM	20 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
Clusters, n (%)									
Cluster 1	0 (0.0)	5 (62.5)	4 (66.7)	3 (42.9)	6 (75.0)	9 (81.8)	1 (33.3)	2 (28.6)	<0.001[c]
Cluster 2	0 (0.0)	3 (37.5)	2 (33.3)	4 (57.1)	2 (25.0)	2 (18.2)	2 (66.7)	5 (71.4)	
Cluster 3	20 (100.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	
Maternal HOMA-IR in the 1st week postpartum									
Elevated (> 2.0), n (%)	1 (5.0)	1 (12.5)	1 (16.7)	1 (14.3)	1 (12.5)	4 (36.4)	0 (0.0)	0 (0.0)	0.324[c]
Non-elevated (≤ 2.0), n (%)	19 (15.0)	7 (87.5)	5 (83.3)	6 (85.7)	7 (87.5)	7 (63.6)	3 (100.0)	7 (100.0)	

Table 5. Comparison of glucose phenotypes across study groups – cont.

Selected gestational and perinatal outcomes	Glucose phenotypes based on OGTT								p-value
	normal OGTT (1)	elevated FGL + normal 1-h POGL + normal 2-h POGL (2)	normal FGL + elevated 1-h POGL + normal 2-h POGL (3)	normal FGL + normal 1-h POGL + elevated 2-h POGL (4)	elevated FGL + elevated 1-h POGL + elevated 2-h POGL (5)	elevated FGL + elevated 1-h POGL + normal 2-h POGL (6)	elevated FGL + normal 1h POGL + elevated 2-h POGL (7)	normal FGL + elevated 1h POGL + elevated 2-h POGL (8)	
	n = 20	n = 8	n = 6	n = 7	n = 8	n = 11	n = 3	n = 7	
Maternal QUICKI in the 1st week postpartum									
Decreased (<0.34), n (%)	1 (5.0)	1 (12.5)	1 (16.7)	1 (14.3)	1 (12.5)	4 (36.4)	0 (0.0)	0 (0.0)	0.324[c]
Normal (≥0.34), n (%)	19 (15.0)	7 (87.5)	5 (83.3)	6 (85.7)	7 (87.5)	7 (63.6)	3 (100.0)	7 (100.0)	
Maternal HbA1c [%] in the 1st week postpartum									
Elevated (≥6.1%), n (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (12.5)	1 (9.1)	0 (0.0)	0 (0.0)	0.573[c]
Non-elevated (<6.1%), n (%)	20 (100.0)	8 (100.0)	6 (100.0)	7 (100.0)	7 (87.5)	10 (90.9)	3 (100.0)	7 (100.0)	
Maternal classification by pre-gestational BMI, n (%)									
Normal	17 (85.0)	2 (25.0)	2 (33.3)	4 (57.1)	2 (25.0)	5 (45.4)	1 (33.3)	7 (100.0)	0.014[c]
Overweight	1 (5.0)	4 (50.0)	3 (50.0)	1 (14.3)	2 (25.0)	3 (27.3)	2 (66.7)	0 (0.0)	
Obese	2 (10.0)	2 (25.0)	1 (16.7)	2 (28.6)	4 (50.0)	3 (27.3)	0 (0.0)	0 (0.0)	
Maternal weight gain during pregnancy in reference to pre-gestational BMI, n (%)									
Below recommendations	2 (10.0)	1 (12.5)	2 (33.3)	2 (28.6)	3 (37.5)	3 (27.3)	2 (66.7)	3 (42.9)	0.018[c]
Within recommendations	5 (25.0)	7 (87.5)	2 (33.3)	4 (57.1)	0 (0.0)	5 (45.4)	1 (33.3)	2 (28.6)	
Above recommendations	13 (65.0)	0 (0.0)	2 (33.3)	1 (14.3)	5 (62.5)	3 (27.3)	0 (0.0)	2 (28.6)	
Nicotinism before pregnancy, n (%)	5 (25.0)	3 (37.5)	2 (33.3)	3 (42.9)	4 (50.0)	3 (27.3)	1 (33.3)	1 (14.3)	0.858[c]
Maternal hypothyroidism, n (%)									
Chronic (onset before the pregnancy)	2 (10.0)	3 (37.5)	0 (0.0)	2 (28.6)	4 (50.0)	6 (54.5)	1 (33.3)	2 (28.6)	0.335[c]
Gestational (onset during the pregnancy)	4 (20.0)	1 (12.5)	3 (50.0)	4 (57.1)	1 (12.5)	1 (9.1)	0 (0.0)	2 (28.6)	
None	14 (70.0)	4 (50.0)	3 (50.0)	1 (14.3)	3 (37.5)	4 (36.4)	2 (66.7)	3 (42.9)	
Maternal hypertension, n (%)									
Chronic (onset before the pregnancy)	0 (0.0)	1 (12.5)	1 (16.7)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (14.3)	0.422[c]
Pregnancy induced	1 (5.0)	1 (12.5)	1 (16.7)	0 (0.0)	1 (12.5)	3 (27.3)	1 (33.3)	0 (0.0)	
None	19 (15.0)	6 (75.0)	4 (66.7)	7 (100.0)	7 (87.5)	8 (72.7)	2 (66.7)	6 (85.7)	
VI, n (%)	1 (5.0)	2 (25.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (9.1)	1 (33.3)	1 (14.3)	0.361[c]
UTI, n (%)	0 (0.0)	0 (0.0)	1 (16.7)	1 (14.3)	2 (25.0)	1 (9.1)	0 (0.0)	1 (14.3)	0.456[c]
SGA, n (%)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	0 (0.0)	1 (9.1)	1 (33.3)	1 (14.3)	0.153[c]
LGA, n (%)	4 (20.0)	2 (25.0)	1 (16.7)	0 (0.0)	3 (37.5)	1 (9.1)	0 (0.0)	1 (14.3)	0.610[c]
Neonatal respiratory distress after birth, n (%)	1 (5.0)	0 (0.0)	0 (0.0)	1 (14.3)	0 (0.0)	2 (18.2)	2 (66.7)	0 (0.0)	0.014[c]
Congenital malformation in neonate, n (%)	0 (0.0)	1 (12.5)	1 (16.7)	1 (14.3)	0 (0.0)	0 (0.0)	0 (0.0)	1 (14.3)	0.486[c]
Method of feeding of a newborn, n (%)									
Exclusive breastfeeding	15 (75.0)	4 (50.0)	0 (0.0)	3 (42.9)	4 (50.0)	2 (18.2)	2 (66.7)	3 (42.9)	0.026[c]
Predominant breastfeeding	5 (25.0)	4 (50.0)	6 (100.0)	4 (57.1)	4 (50.0)	9 (81.8)	1 (33.3)	4 (57.1)	

[c] – χ^2 test; OGTT – oral glucose test; FGL – fasting glucose level; 1-h POGL – 1 h post OGTT glucose level; 2-h POGL – 2 h post OGTT glucose level; n – number of patients; non-GDM – control group; GDM G1 – group of mothers with GDM who were administered diet and physical activity; GDM G2 – group of mothers with GDM receiving insulin treatment; BMI – body mass index; HbA1c – hemoglobin A1c; HOMA-IR – Homeostatic Model Assessment–Insulin Resistance; QUICKI – Quantitative Insulin Sensitivity Check Index; SGA – small for gestational age; LGA – large for gestational age; VI – vaginal infections; UTI – urinary tract infections.

Discussion

The average number of births per year in the authors' unit is 2,000. Of these, newborns with a gestational age greater than 35 +0/7 weeks account for approx. 1,700 (85%). Accordingly, 50 GDM mothers in the current study represent 2.9% of all women who gave birth in the authors' unit. This result is similar to the reported GDM prevalence of 3.4% among pregnant women in Poland.

The most effective way to manage GDM and reduce the risk of complications is early detection and appropriate treatment. Diet, physical activity, and insulin therapy are the available therapeutic options. Reviews on the optimal management of GDM state that if diet and physical activity alone are ineffective in achieving glycemic control within 2 weeks, insulin should be administered in conjunction with dietary modifications and physical activity.[24] Accordingly, the type of therapeutic interventions necessary to achieve normal glucose levels corresponds to the extent of gestational glucose metabolism disorders and, consequently, IR.[24,25] Insulin resistance combined with impaired insulin secretion is the primary underlying cause of GDM. It is difficult to determine the degree of IR and insulin secretion in women with newly diagnosed GDM. Predicting insulin requirements during pregnancy is even more problematic because IR increases gradually over the course of pregnancy due to hormonal changes. The process of developing IR in pregnancy occurs in all pregnant women. Available studies have shown a decrease in insulin sensitivity of almost 60% during a normal pregnancy. However, GDM is more likely to develop in women with additional risk factors (genetic, environmental, etc.). Gestational diabetes mellitus is considered a temporary condition and is usually resolved postpartum, within 6–12 weeks of the puerperium.[26,27]

The authors, as members of the neonatal team, were the 1st to have the opportunity to recruit the mothers and assess the course of the pregnancies in the postpartum period. The evaluation included the effectiveness of glucose control, the results of laboratory tests, and a medical interview regarding maternal comorbidities. However, no results for IR indicators at various stages of pregnancy were available. We therefore assumed that GDM mothers treated with insulin would have a higher degree of glucose intolerance compared with those treated with diet and physical activity alone. We also aimed to compare whether these 2 groups differed in IR marker levels during the 1st week after childbirth, when other analyses were planned.

This assumption is consistent with the results of our study. Mothers with GDM were found to have higher levels of HOMA-IR, QUICKI, and HbA1c in the 1st week postpartum, as well as higher glucose levels at 3 time points of the OGTT than non-GDM mothers. In addition, the GDM G2 group (treated with insulin therapy) had higher values than the GDM G1 group (dietary changes and physical activity only), except for the 1 h OGTT. As only one of 3 criteria needs to be met for a diagnosis of GDM, the GDM-affected study groups differed in the incidence of abnormal glucose screening at 3 time points of the OGTT. The highest incidence of abnormal glucose levels at each time point was found in the GDM G2 group. The same observation applies to 1st-trimester fasting glucose levels. However, the incidence of selected complications did not differ between the study groups.

The results regarding breastfeeding rates warrant particular attention. The results documented an unexpectedly limited prevalence of exclusive breastfeeding among mothers with GDM, particularly among those classified as GDM G2. Women affected by GDM exhibited less favorable breastfeeding outcomes compared to those without GDM, despite comparable rates of breastfeeding initiation. According to the available literature, there appears to be a heightened propensity for delayed onset of lactogenesis II in this population.[26] This is further compounded by the fact that women with GDM report greater difficulty in producing sufficient milk compared with women without diabetes. Consequently, infants of mothers with GDM are introduced to human milk substitutes (e.g., formula) at an earlier age than infants of women without a history of diabetes. The use of formula is also associated with earlier cessation of breastfeeding.[27–31]

The trend toward formula-supplemented feeding among GDM mothers is clinically significant: formula feeding in GDM-exposed infants is linked to a higher risk of obesity during early childhood, whereas exclusive breastfeeding offers protective effects.[28,31] This highlights an actionable intervention – supporting and encouraging breastfeeding in GDM contexts may help mitigate metabolic risks in offspring.

Several surrogate measures have been introduced to evaluate insulin sensitivity, including HOMA-IR and QUICKI. These measures, along with HbA1c, have also been considered biomarkers capable of predicting adverse pregnancy outcomes in women with GDM.[25,32] However, the threshold of HOMA-IR for the assessment of IR has not been universally agreed upon. This is due to various factors correlated with HOMA-IR, including age, sex, ethnicity, and body weight. HOMA-IR demonstrates a robust correlation with BMI, as evidenced by substantial differences in HOMA-IR between individuals with normal weight and those categorized as obese.[32]

The present study demonstrated that HOMA-IR exhibited an upward trend in conjunction with elevated pre-pregnancy BMI, with the highest levels observed in the GDM G2 group. In contrast, HOMA-IR > 2.0 and QUICKI < 0.34 are most often considered indicative of IR.[15,16] These levels were found in 9 (12.9%) mothers in the 1st week postpartum, with a predominance among insulin-treated GDM mothers. The authors suggest that a higher degree of glucose intolerance during pregnancy requiring insulin administration may contribute to the development of persistent IR in the early postpartum period.

Based on international studies, 3rd-trimester HbA1c levels have been identified as a determinant of LGA in pregnancies affected by GDM.[33] However, early control of GDM (before 34 weeks) was found to result in an 18% lower rate of LGA infants compared with late control of GDM (after 34 weeks).[34] In the present study, all patients diagnosed with GDM were provided with dietary and lifestyle recommendations, as well as self-monitoring of blood glucose levels. In addition, stricter weight gain targets during pregnancy were imposed on participants with elevated BMI values before pregnancy.

In a large proportion of mothers, weight gain during pregnancy was classified as above recommendations, with the highest incidence observed in the non-GDM group. Therefore, it was not possible to determine whether the nonsignificant difference in the incidence of LGA between women with higher BMI and a higher prevalence of overweight and obesity (GDM G2) was attributable to lifestyle treatment or to stricter weight gain targets. On the other hand, non-GDM mothers who did not receive lifestyle intervention during pregnancy had the highest gestational weight gain, which did not result in an increased incidence of LGA. Similar results were reported by Sun et al.[35] The aforementioned study also emphasized the importance of maintaining pre-pregnancy BMI within the optimal range and preventing excessive weight gain to reduce excessive IR in the 2nd trimester.

Despite evidence indicating the potential of HbA1c as a biomarker for predicting adverse outcomes in pregnant women with GDM, it is important to consider maternal age, pre-pregnancy BMI, and gestational weight gain when assessing the relationship between HbA1c and adverse perinatal outcomes.[26] Maternal age, BMI, and excessive weight gain prior to 24 weeks' gestation have been recognized as contributing factors to the development of GDM and IR. A comprehensive analysis of the Taiwanese population revealed an elevated risk of IR in individuals with high BMI values. Consistent with the findings of Sun et al., increased pre-pregnancy BMI and weight gain prior to the diagnosis of GDM were strongly associated with an increased risk of impaired glucose tolerance in the 2nd trimester.[35] A total of 293 women with a history of GDM or a pre-pregnancy BMI greater than 30 kg/m^2 were enrolled in the RADIEL trial and subsequently divided into 2 groups. Patients in the intervention group received early individualized counseling on diet, physical activity, and weight control prior to 20 weeks' gestation. The incidence of GDM in this group was significantly lower than in the control group.[36]

Therefore, our results confirm that appropriate antenatal care, health education, and rigorous counseling are effective in preventing, or at least reducing, the risk of complications in women with GDM and their newborns.

The implementation of cluster analysis in the present study enabled a distinct categorization of patients based not only on allocation to the study groups but also on pregestational BMI and gestational weight gain. These factors are recognized as key determinants of IR and adverse pregnancy outcomes. The mean M–S distance between the clusters and the original groups was 0.027, indicating a high degree of agreement between the 2 categorizations. The observed discrepancies were confined to a single case, which was classified differently by the clustering algorithm than in the original grouping. An analysis of the patient's characteristics suggests that the patient's profile may have been at the threshold between groups, thereby explaining the assignment to a different cluster. This finding may indicate that the clustering captured important differences in the data that were not fully reflected in the initial classification. Consequently, this approach may represent a valuable addition to, or potential refinement of, the original diagnostic grouping.

Given the observed similarity in neonatal birth weight, mean birth weight, and birth weight percentile across both study groups and clusters, the findings related to neonatal body composition were likewise consistent across these groups. This finding supports the effectiveness of glycemic and weight control measures in mitigating macrosomia.

Identification of maternal glucose phenotypes confirmed that the degree of glucose metabolism disorders varied among mothers. This variation may have been influenced by other factors affecting overall maternal health (e.g., hypothyroidism, obesity, and nicotine dependence) as well as the course of pregnancy (e.g., gestational weight gain and urinary tract infection). As a result, pregnancy outcomes may also have been affected, including neonatal LGA or respiratory distress, success of exclusive breastfeeding, and selected IR marker levels in the early postpartum period.

Furthermore, this analysis suggests a potential relationship between 2 factors –maternal pre-pregnancy BMI and excessive gestational weight gain – on the one hand, and adverse pregnancy outcomes, on the other. However, the small number of mothers within each glucose phenotype limits the ability to draw definitive conclusions regarding associations between specific glucose phenotypes and selected pregnancy complications. It is also inappropriate to generalize these results to the Polish population.

Nevertheless, the application of cluster analysis and glucose phenotype stratification in mothers with GDM facilitates the identification of high-risk groups, thereby enabling targeted counseling and monitoring aimed at preventing future metabolic disorders.

The findings of the present study did not demonstrate a significant difference in IR marker outcomes between mothers with newly diagnosed GDM and those with a history of GDM. However, short-term studies have shown that IR remains higher in women with a history of GDM.[37,38] This inconsistency may be attributable to the small number of mothers with previous GDM included in our study.

By identifying high-risk groups using these markers, healthcare providers may implement personalized monitoring and intervention strategies to improve maternal

and fetal outcomes by addressing subgroup-specific risk profiles.

The classification of GDM patients at high risk of complications depends on understanding the sociodemographic and health factors that inform these risk assessments. A recent Polish study found that non-insulin-based indices of IR – the triglyceride and glucose (TyG) index and the metabolic score for insulin resistance (METS-IR) – were significantly associated with older age, smoking, and elevated systolic blood pressure, indicating coexisting cardiometabolic risks. Furthermore, an elevated TyG index was associated with higher BMI and alcohol consumption. Conversely, lower educational attainment was associated with higher METS-IR scores.

In patients with GDM, who already experience increased IR, these findings underscore the importance of incorporating sociodemographic metrics (e.g., age, education level, smoking, alcohol consumption) and key health indicators (e.g., BMI and blood pressure) into risk-stratification models.[39] This multidimensional approach may help guide early interventions, personalized monitoring, and tailored support strategies aimed at mitigating both metabolic and obstetric complications.

The concept of individualized medicine has been proposed as a catalyst for advancements in healthcare, with the capacity to enhance the effectiveness, efficiency, and patient-centeredness of healthcare services. By focusing on the unique characteristics of each patient, this approach aims to improve outcomes and reduce unnecessary treatments. A substantial body of clinical research has demonstrated that the personalization of nutritional plans to meet the needs of individual patients – taking into account factors such as body weight, metabolic status, energy demands, and meal timing (a concept referred to as "chrononutrition") – can lead to meaningful improvements in patient outcomes. Additionally, reduced gestational weight gain has been shown to be associated with favorable pregnancy outcomes.[40,41]

Limitations of the study

The primary limitation of this study is the limited sample size. While the preliminary results provide a general overview of the issue addressed, these findings require verification through additional research involving a larger cohort. Moreover, the reproducibility of these results cannot be assured. One of the key concerns relates to the potential inaccuracy of assumptions based on the available data.

A further limitation is the absence of IR marker assessments during pregnancy, which limits the ability to compare results across different stages and precludes conclusions regarding potential changes in IR over subsequent weeks. An exception is HbA1c, as the average lifespan of erythrocytes is 60–90 days. Consequently, HbA1c levels provide insight into estimated glucose concentrations during the final 2–3 months of pregnancy.

Conclusions

Markers of IR, such as HbA1c and HOMA-IR, in the 1st week after delivery varied significantly between mothers with GDM and healthy controls, with the highest incidence of abnormal results observed among insulin-treated GDM mothers. The degree of impaired glucose metabolism, along with maternal preconceptional BMI and gestational weight change, had a significant influence on pregnancy outcomes. These variables appear to be closely associated with the maternal glucose profile, thereby influencing maternal OGTT results at various time points during pregnancy.

With respect to most pregnancy outcomes, including anthropometric measurements of newborns and the composition of their body compartments, no substantial differences were observed between the study groups. However, an inverse relationship was identified between the severity of glucose tolerance impairment and maternal weight accumulation during gestation.

Conversely, in mothers with impaired glucose metabolism during pregnancy, abnormal body weight or excessive weight gain may lead to a higher incidence of LGA in newborns. Adequate glycemic control and effective treatment, regardless of the type of intervention, are important factors influencing pregnancy outcomes. Both cluster analysis and stratification of glucose phenotypes among mothers with GDM are valuable methods for identifying distinct patient groups based on a combination of risk factors. These approaches may be helpful in identifying patients at high risk of adverse pregnancy outcomes, as well as mothers and newborns who require careful counseling and monitoring to prevent future metabolic disorders..

The concept of individualized medicine in GDM adopts a multifaceted approach, encompassing personalized nutrition, physical activity regimens, pharmacological therapy selection, digital guidance, and genetic and metabolic profiling. This integrated strategy aims to provide earlier, more effective, and more patient-centered care, thereby enhancing outcomes for both the mother and child.

In summary, the findings of the present study support the view that GDM is not merely a condition confined to pregnancy but rather part of a broader continuum of maternal–fetal metabolic health. These findings support the importance of early intervention, personalized care guided by BMI and glucose profiles, and the promotion of breastfeeding to enhance both immediate and long-term health outcomes for mothers and their children.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.15862275. The package contains the following files:

Supplementary Table 1. Results of the Shapiro–Wilk test for normality and Levene's test for homogeneity of variances across selected characteristics and pregnancy outcomes in clusters and groups.

Supplementary Table 2. Results of regression analysis (generalized linear model).

Data Availability Statement

The participants in this study did not provide written consent for their data to be shared publicly; therefore, due to the sensitive nature of the research, the supporting data are not available.

Consent for publication

Not applicable.

Use of AI and AI-assisted technology

Not applicable.

ORCID iDs

Karolina Karcz https://orcid.org/0000-0002-4271-2626
Paulina Gaweł https://orcid.org/0000-0001-5991-4470
Barbara Królak-Olejnik https://orcid.org/0000-0002-6493-0333

References

1. Caughey AB. Gestational diabetes mellitus: Obstetric issues and management. Alphen aan den Rijn, the Netherlands; Wolters Kluwer; 2024. https://hcp.mn/sites/default/files/2024-01/Gestational_diabetes_mellitus_Obstetric_issues_and_management_en.pdf. Accessed November 22, 2024.
2. International Association of Diabetes and Pregnancy Study Groups Consensus Panel. International Association of Diabetes and Pregnancy Study Groups Recommendations on the Diagnosis and Classification of Hyperglycemia in Pregnancy. *Diabetes Care*. 2010;33(3):676–682. doi:10.2337/dc09-1848
3. Dalfrà MG, Burlina S, Del Vescovo GG, Lapolla A. Genetics and epigenetics: New insight on gestational diabetes mellitus. *Front Endocrinol (Lausanne)*. 2020;11:602477. doi:10.3389/fendo.2020.602477
4. Kampmann U, Knorr S, Fuglsang J, Ovesen P. Determinants of maternal insulin resistance during pregnancy: An updated overview. *J Diabetes Res*. 2019;2019:5320156. doi:10.1155/2019/5320156
5. Guariguata L, Linnenkamp U, Beagley J, Whiting DR, Cho NH. Global estimates of the prevalence of hyperglycaemia in pregnancy. *Diabetes Res Clin Pract*. 2014;103(2):176–185. doi:10.1016/j.diabres.2013.11.003
6. Wender-Ożegowska E, Bomba-Opoń D, Brązert J, et al. Standards of Polish Society of Gynecologists and Obstetricians in management of women with diabetes. *Ginekol Pol*. 2018;89(6):341–350. doi:10.5603/gp.a2018.0059
7. World Health Organization (WHO). Diagnostic Criteria and Classification of Hyperglycaemia First Detected in Pregnancy. Geneva, Switzerland: World Health Organization (WHO); 2013. https://apps.who.int/iris/handle/10665/85975. Accessed November 22, 2024.
8. American Diabetes Association. Classification and Diagnosis of Diabetes: *Standards of Medical Care in Diabetes–2020. Diabetes Care*. 2020;43(Suppl 1):S14–S31. doi:10.2337/dc20-s002
9. Ye W, Luo C, Huang J, Li C, Liu Z, Liu F. Gestational diabetes mellitus and adverse pregnancy outcomes: Systematic review and meta-analysis. *BMJ*. 2022;377:e067946. doi:10.1136/bmj-2021-067946
10. Farahvar S, Walfisch A, Sheiner E. Gestational diabetes risk factors and long-term consequences for both mother and offspring: A literature review. *Exp Rev Endocrinol Metab*. 2019;14(1):63–74. doi:10.1080/17446651.2018.1476135
11. Freeman AM, Acevedo LA, Pennings N. Insulin resistance. In: *StatPearls*. Treasure Island, USA: StatPearls Publishing; 2025:Bookshelf ID: NBK507839. http://www.ncbi.nlm.nih.gov/books/NBK507839. Accessed July 15, 2025.
12. Johnson JD. On the causal relationships between hyperinsulinaemia, insulin resistance, obesity and dysglycaemia in type 2 diabetes. *Diabetologia*. 2021;64(10):2138–2146. doi:10.1007/s00125-021-05505-4
13. Czech MP. Insulin action and resistance in obesity and type 2 diabetes. *Nat Med*. 2017;23(7):804–814. doi:10.1038/nm.4350
14. American College of Obstetricians and Gynecologists (ACOG). Committee Opinion No. 548: Weight Gain During Pregnancy. *Obstet Gynecol*. 2013;121(1):210–212. doi:10.1097/01.aog.0000425668.87506.4c
15. Singh B. Surrogate markers of insulin resistance: A review. *World J Diabetes*. 2010;1(2):36. doi:10.4239/wjd.v1.i2.36
16. Placzkowska S, Pawlik-Sobecka L, Kokot I, Piwowar A. Indirect insulin resistance detection: Current clinical trends and laboratory limitations. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2019;163(3):187–199. doi:10.5507/bp.2019.021
17. Yeung RO, Retnakaran R, Savu A, Butalia S, Kaul P. Gestational diabetes: One size does not fit all. An observational study of maternal and neonatal outcomes by maternal glucose profile. *Diabet Med*. 2024;41(2):e15205. doi:10.1111/dme.15205
18. Marra M, Sammarco R, De Lorenzo A, et al. Assessment of body composition in health and disease using bioelectrical impedance analysis (BIA) and dual energy X-ray absorptiometry (DXA): A critical overview. *Contrast Media Mol Imaging*. 2019;2019:3548284. doi:10.1155/2019/3548284
19. Karcz K, Czosnykowska-Lukacka M, Krolak-Olejnik B. Impact of gestational diabetes and other maternal factors on neonatal body composition in the first week of life: A case-control study. *Ginekol Pol*. 2023;94(2):119–128. doi:10.5603/gp.a2021.0249
20. Maechler M. Package 'cluster'. Cluster Analysis, extended original from Peter Rousseeuw, Anja Struyf and Mia Hubert, version 1.14.3. Comprehensive R Archive Network (CRAN); 2012. http://cran.r-project.org/web/packages/cluster/index.html. Accessed November 22, 2024.
21. Fraley C, Raftery A, Scrucca L. Normal mixture modeling for model-based clustering, classification, and density estimation, version 4.0. Comprehensive R Archive Network (CRAN); 2012. http://cran.r-project.org/web/packages/mclust/index.html. Accessed November 22, 2024.
22. Marczewski E, Steinhaus H. On a certain distance of sets and the corresponding distance of functions. *Colloq Math*. 1958;6:319–327. http://matwbn.icm.edu.pl/ksiazki/cm/cm6/cm6141.pdf.
23. Zhai CX. A note on the expectation-maximization (EM) algorithm. Chicago, USA: Department of Computer Science, University of Illinois at Urbana-Champaign; 2007. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=8c3e7708aa78ba89ce85df96f17de3125def7d4c. Accessed November 22, 2024.
24. Ghamri K, Alsulami S, Alotaibi L, Salem I, Tash R, Yousof S. Determinants of insulin therapy among women with gestational diabetes mellitus: A cross-sectional study. *Niger J Clin Pract*. 2023;26(4):417–423. doi:10.4103/njcp.njcp_447_22
25. Duo Y, Song S, Zhang Y, et al. Predictability of HOMA-IR for gestational diabetes mellitus in early pregnancy based on different first trimester BMI values. *J Pers Med*. 2022;13(1):60. doi:10.3390/jpm13010060
26. Lin J, Jin H, Chen L. Associations between insulin resistance and adverse pregnancy outcomes in women with gestational diabetes mellitus: A retrospective study. *BMC Pregnancy Childbirth*. 2021;21(1):526. doi:10.1186/s12884-021-04006-x
27. Sharma AK, Singh S, Singh H, et al. Deep insight of the pathophysiology of gestational diabetes mellitus. *Cells*. 2022;11(17):2672. doi:10.3390/cells11172672
28. Doughty KN, Taylor SN. Barriers and benefits to breastfeeding with gestational diabetes. *Semin Perinatol*. 2021;45(2):151385. doi:10.1016/j.semperi.2020.151385
29. Haile ZT, Oza-Frank R, Azulay Chertok IR, Passen N. Association between history of gestational diabetes and exclusive breastfeeding at hospital discharge. *J Hum Lact*. 2016;32(3):NP36–NP43. doi:10.1177/0890334415618936
30. Nguyen PTH, Binns CW, Nguyen CL, et al. Gestational diabetes mellitus reduces breastfeeding duration: A prospective cohort study. *Breastfeed Med*. 2019;14(1):39–45. doi:10.1089/bfm.2018.0112

31. Manerkar K, Harding J, Conlon C, McKinlay C. Maternal gestational diabetes and infant feeding, nutrition and growth: A systematic review and meta-analysis. *Br J Nutr.* 2020;123(11):1201–1215. doi:10.1017/s0007114520000264
32. Muhuza MPU, Zhang L, Wu Q, Qi L, Chen D, Liang Z. The association between maternal HbA1c and adverse outcomes in gestational diabetes. *Front Endocrinol (Lausanne).* 2023;14:1105899. doi:10.3389/fendo.2023.1105899
33. Fonseca L, Saraiva M, Amado A, et al. Third trimester HbA1c and the association with large-for-gestational-age neonates in women with gestational diabetes. *Arch Endocrinol Metab*. 2021;65(3):328–335. doi:10.20945/2359-3997000000366
34. Shushan A, Ezra Y, Samueloff A. Early treatment of gestational diabetes reduces the rate of fetal macrosomia. *Am J Perinatol*. 1997; 14(5):253–256. doi:10.1055/s-2007-994138
35. Sun Y, Juan J, Xu Q, Su R, Hirst JE, Yang H. Increasing insulin resistance predicts adverse pregnancy outcomes in women with gestational diabetes mellitus. *J Diabetes*. 2020;12(6):438–446. doi:10.1111/1753-0407.13013
36. Koivusalo SB, Rönö K, Klemetti MM, et al. Gestational diabetes mellitus can be prevented by lifestyle intervention: The Finnish Gestational Diabetes Prevention Study (RADIEL). *Diabetes Care*. 2016;39(1):24–30. doi:10.2337/dc15-0511
37. Miao Z, Wu H, Ren L, et al. Long-term postpartum outcomes of insulin resistance and β-cell function in women with previous gestational diabetes mellitus. *Int J Endocrinol*. 2020;2020:7417356. doi:10.1155/2020/7417356
38. Homko C, Sivan E, Chen X, Reece EA, Boden G. Insulin secretion during and after pregnancy in patients with gestational diabetes mellitus. *J Clin Endocrinol Metab*. 2001;86(2):568–573. doi:10.1210/jcem.86.2.7137
39. Polak M, Nowicki GJ, Chrzanowska-Wąsik M, Ślusarska BJ. Do sociodemographic and health predictors affect the non-insulin-based insulin resistance index? A cross-sectional study [published online as ahead of print on November 12, 2024]. *Adv Clin Exp Med*. 2024. doi:10.17219/acem/191200
40. Zhang X, Wu Y, Miao L. Study on the effects of individualized nutritional intervention on pregnancy outcome and neonatal immune function in patients with gestational diabetes mellitus. *Biomed Res Int*. 2022;2022(1):3246784. doi:10.1155/2022/3246784
41. Luo JY, Chen LG, Yan M, Mei YJ, Cui YQ, Jiang M. Effect of individualized nutrition interventions on clinical outcomes of pregnant women with gestational diabetes mellitus. *World J Diabetes*. 2023;14(10): 1524–1531. doi:10.4239/wjd.v14.i10.1524

Impact of patient–physician communication and disease knowledge on treatment adherence in glaucoma patients

Katarzyna Malewicz[1,A,D,F], Mariusz Chabowski[2,3,C,E,F], Jakub Staś[4,A,C,D], Anna Maria Cybulska[5,C,E,F], Anna Szymańska-Chabowska[6,B,D,F], Beata Jankowska-Polańska[7,8,A,B,E,F]

[1] Faculty of Nursing and Midwifery, Wroclaw Medical University, Poland
[2] 4th Military Teaching Hospital, Wrocław, Poland
[3] Department of Clinical Surgical Sciences, Faculty of Medicine, Wroclaw University of Science and Technology, Poland
[4] Student Research Group No. 180, Faculty of Medicine, Wroclaw Medical University, Poland
[5] Department of Nursing, Pomeranian Medical University in Szczecin, Poland
[6] Department of Diabetology, Hypertension and Internal Medicine, Faculty of Medicine, Wroclaw Medical University, Poland
[7] Department of Preclinical Sciences, Pharmacology and Medical Diagnostics, Faculty of Medicine, Wroclaw University of Science and Technology, Poland
[8] Center for Research and Innovation, 4th Military Clinical Hospital, Wrocław, 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(5):795–808

Address for correspondence
Mariusz Chabowski
E-mail: mariusz.chabowski@gmail.com

Funding sources
None declared

Conflict of interest
None declared

Received on May 28, 2025
Reviewed on June 20, 2025
Accepted on July 17, 2025

Published online on February 26, 2026

Cite as
Malewicz K, Chabowski M, Staś J, Cybulska AM, Szymańska-Chabowska A, Jankowska-Polańska B. Impact of patient–physician communication and disease knowledge on treatment adherence in glaucoma patients. *Adv Clin Exp Med*. 2026;35(5):795–808. doi:10.17219/acem/208352

DOI
10.17219/acem/208352

Abstract

Background. Glaucoma is a chronic, progressive optic neuropathy that can lead to irreversible blindness if left untreated. It is the leading cause of irreversible blindness globally. Effective management relies heavily on the control of intraocular pressure (IOP), typically through lifelong pharmacological treatment, but adherence to therapy is often a challenge due to the asymptomatic nature of the disease.

Objectives. The primary aim of this study was to evaluate the predictive influence of the doctor–patient relationship, the quality of communication between doctor and patient, patient knowledge about glaucoma, and quality of life when adhering to therapeutic recommendations among a group of Polish glaucoma patients.

Materials and methods. This study was conducted at the Ophthalmology Outpatient Clinic of the University Clinical Hospital in Wrocław, Poland. A total of 190 patients were enrolled, and adherence to treatment was assessed using the Adherence to Refills and Medications Scale (ARMS). Additional variables included patient–physician communication (CAT-14), relationship quality (PDRQ-9), visual function (NEI VFQ-25), and glaucoma knowledge (GKQ-10).

Results. The results indicated that 58.9% of patients showed low adherence to their prescribed treatment. Satisfaction with doctor–patient communication, higher levels of knowledge about glaucoma, and better visual function were significantly associated with better adherence. The quality of the patient–physician relationship, while not statistically significant, also showed potential positive effects on adherence.

Conclusions. Improving patient education and enhancing the quality of communication between healthcare providers and patients are critical to increasing adherence to glaucoma treatment. Importantly, the study provides a valuable evidence base for developing adherence-enhancing strategies within the Polish healthcare system. Given the scarcity of such data in Central and Eastern Europe, our results may also inform similar efforts in the region. Designing culturally sensitive and system-specific interventions, such as brief communication training for ophthalmologists or structured patient education programs, could help address persistent adherence challenges in chronic ophthalmic care.

Key words: glaucoma, communication, patient education, physician–patient relations, medication adherence

Highlights

- Low treatment adherence: A substantial proportion (58.9%) of Polish glaucoma patients showed low adherence to prescribed therapy.
- Communication and knowledge: Improved doctor–patient communication and greater patient knowledge about glaucoma were significantly associated with better adherence.
- Role of visual function: Patients with better visual function were more likely to adhere to glaucoma treatment.
- Patient–physician relationship: Although not statistically significant, a stronger patient–physician relationship showed a positive trend toward improved adherence, suggesting potential value for future interventions.

Background

Glaucoma is a chronic, progressive optic neuropathy characterized by damage to the optic nerve and the retinal nerve fiber layer, which may lead to irreversible blindness if left untreated.[1] Globally, glaucoma is the leading cause of irreversible blindness, affecting more than 76 million individuals, with projections indicating that this number will increase to 112 million by 2040.[2] It is particularly prevalent in older populations, non-White ethnicities, and individuals with a family history of the disease.[1,3]

There are 2 main types of glaucoma: primary open-angle glaucoma (OAG) and angle-closure glaucoma (ACG). The former is the most common type and is often asymptomatic until significant vision loss occurs, with up to 50% of patients unaware of their condition.[4] In contrast, the latter displays more acute symptoms, such as eye pain, nausea, and blurred vision. If left untreated, it can lead to rapid and irreversible vision loss.[1]

Given its progressive nature and the potential for irreversible vision loss, early diagnosis and intervention are critical for the management of glaucoma. Diagnostic methods include the measurement of intraocular pressure (IOP), visual field testing, and imaging techniques such as optical coherence tomography (OCT). Intraocular pressure remains the only modifiable risk factor for glaucoma, making its control the primary focus of treatment.[1,5,6]

Pharmacological treatment, particularly in the form of eye drops, plays a crucial role in slowing the progression of glaucoma. Medications such as prostaglandin analogues, β-blockers, and carbonic anhydrase inhibitors act to reduce the production or improve the outflow of aqueous humor, thereby lowering IOP. Patients require lifelong treatment and regular follow-up to ensure effective IOP control and to monitor disease progression. Failure to adhere to prescribed therapy can result in inadequate IOP control, which significantly increases the risk of disease progression and permanent visual impairment.[1,7,8]

Adherence to treatment is defined as the extent to which a patient's behavior in taking medications corresponds to the prescribed treatment regimen.[9] Several factors contribute to difficulties in maintaining adherence to glaucoma medications, particularly given the chronic and often asymptomatic nature of the disease. Without noticeable symptoms, patients may not fully recognize the severity of the condition or the importance of consistent medication use. This lack of symptomatic feedback can result in reduced adherence, as patients may mistakenly believe that the medication is unnecessary or provides no benefit.[10,11] Adherence to treatment recommendations is a complex, multifaceted phenomenon. Various factors influencing adherence have been widely discussed in the literature.[10–12] A common conclusion of many studies on adherence is the need for continued research and deeper analysis of the determinants of adherence.[13]

Adherence to treatment recommendations, particularly in chronic ophthalmic conditions such as glaucoma, dry eye disease (DED), and age-related macular degeneration (AMD), remains a global challenge. Numerous international studies have demonstrated that non-adherence significantly impairs treatment effectiveness and increases the risk of disease progression and vision loss. For example, in a large-scale Japanese study on DED, only 10.2% of patients used eye drops at the frequency recommended in the treatment guidelines, with most patients applying drops only when experiencing symptoms, highlighting a symptom-driven, inconsistent use pattern.[14] In Brazil, electronic monitoring of glaucoma medication use demonstrated that 28% of patients were non-adherent, and this subgroup had significantly higher rates of disease progression and loss to follow-up rates.[15] Likewise, an Irish study of patients with intermediate or advanced AMD found that 40% of patients eligible for AREDS2 nutritional supplementation were not using the recommended vitamins, often due to a lack of physician recommendation or concerns about cost.[16]

These findings highlight the multifactorial nature of adherence and the influence of both systemic and individual-level determinants, including treatment complexity, health literacy, symptom visibility, and healthcare-related communication. While some barriers are universal, others are shaped by local healthcare systems, cultural expectations, and patient–physician dynamics. In Poland, adherence in ophthalmology has received limited empirical attention, despite an aging population and a high prevalence of chronic eye diseases. Structural aspects of the Polish

healthcare system, such as the organization of outpatient care, reimbursement policies, and variability in access to ophthalmologists, may pose distinct challenges to maintaining long-term adherence. Accordingly, investigating predictors of adherence in a Polish glaucoma cohort not only addresses a notable research gap but may also inform culturally and system-level tailored interventions.

Objectives

The aim of this study was to assess the predictive value of the doctor–patient relationship, the quality of patient–physician communication, patient knowledge about glaucoma, and quality of life in relation to adherence to the recommended treatment in a cohort of Polish patients receiving pharmacological treatment for glaucoma.

Materials and methods

The study was conducted between January and September 2019 at the Ophthalmology Outpatient Clinic, University Clinical Hospital, in Wrocław, Poland. A total of 190 patients who met the inclusion criteria were enrolled. Each participant received written information detailing the study's purpose and procedures, and written informed consent was obtained prior to participation. The study was conducted during a single follow-up visit at the ophthalmology clinic. The research was approved by the Bioethics Committee of Wroclaw Medical University (approval No. KB 305/2018).

Inclusion criteria included a confirmed diagnosis of glaucoma at least 6 months earlier, current use of anti-glaucoma pharmacotherapy, being aged 18 or older, the ability to independently complete the study questionnaires, and a willingness to provide informed consent for participation.

Exclusion criteria comprised a cognitive impairment score of less than 24, which was defined as a Mini-Mental State Examination (MMSE), a lack of the manual dexterity required for the proper administration of anti-glaucoma medications (necessitating assistance from a third party), and the presence of significantly exacerbated chronic conditions.

The Adherence to Refills and Medications Scale (ARMS) is a 12-item questionnaire designed to assess adherence to pharmacological treatment in patients with chronic conditions. Responses are evaluated on a 4-point Likert scale, with options ranging from 1 (never) to 4 (most of the time). The total score ranges from 12, indicating optimal adherence, to 48, reflecting a complete lack of adherence to prescribed pharmacotherapy. This instrument, developed by Kripalani et al. in 2009, specifically targets the assessment of adherence to medication regimens for chronic diseases that require prolonged, often lifelong, treatment. The questionnaire prompts patients to evaluate not only their consistency in taking medications and their adherence to the prescribed dosage and timing, but also their fulfillment of prescriptions and ability to plan and maintain the necessary medication supplies to ensure uninterrupted therapy.[17] The tool has been linguistically validated for use in Polish, with the adaptation conducted among a group of patients with hypertension.[18] For the purposes of the study, an evaluation of the psychometric properties of the assessment tool was conducted. The findings confirmed the tool's validity for use among patients with glaucoma (Cronbach's $\alpha = 0.726$, mean inter-item correlation $r = 0.187$). Patients were categorized into 2 groups based on their adherence levels to therapeutic recommendations: group I – low adherence (16–48 points) and group II – high adherence (12–15 points). This division was determined by the median ARMS score observed in the study group (Me = 16).

The Communication Assessment Tool (CAT-14) was employed to evaluate the quality of communication between patients and physicians. Developed by Makoul et al. in 2007, this instrument consists of 14 items, addressing both verbal and non-verbal communication, as well as the physician's attentiveness, engagement, and empathy. Responses are measured using a 5-point Likert scale: 1 – inadequate, 2 – sufficient, 3 – good, 4 – very good, and 5 – excellent. To assess the physician's interpersonal and communication skills, the average score across all 14 items and the percentage of responses rated as excellent are utilized. The total score ranges from 14 to 70 points, with higher scores indicating a better perceived quality of communication.[19] The questionnaire has been culturally adapted for use in the Polish population.[20]

The Patient–Doctor Relationship Questionnaire (PDRQ-9) was used to assess the quality of the relationship between patients and their physicians. This instrument consists of nine items and was developed by Van der Feltz-Cornelis et al. in 2004. Responses are rated on a Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). The total score ranges from 9, indicating the poorest possible relationship, to 45, representing an ideal relationship.[21–23] As the PDRQ-9 had not undergone cultural adaptation for use in Poland, a validation of the tool was conducted specifically for glaucoma patients as part of this study, demonstrating strong psychometric properties (Cronbach's $\alpha = 0.937$, with an average inter-item correlation of $r = 0.626$).

The National Eye Institute Visual Function Questionnaire (NEI VFQ-25), published by Mangione et al. in 2001, was employed to evaluate the impact of visual impairment and related symptoms on general health domains such as emotional wellbeing, social functioning, and task-related visual capabilities in daily life. The questionnaire includes items that assess general health and vision status, difficulties with specific tasks, and limitations caused by visual problems. The survey consists of 3 sections, with respondents answering a total of 25 closed questions.

The original numeric values from the survey were recoded so that higher scores indicate better functioning. Each item was then converted to a scale from 0 to 100, where 0 represents the lowest possible score and 100 the highest. The final score is calculated as the average of all subscale items completed by the respondent, with lower scores indicating a greater impact of visual impairment on quality of life.[24,25]

The Glaucoma Knowledge Questionnaire (GKQ-10) is a custom-developed tool designed to assess patients' knowledge of glaucoma. The reliability (internal consistency) of the questionnaire was evaluated as part of this study. The questionnaire consists of 10 items assessing sources of knowledge about the disease, its definition, the possibility of a complete cure, the consequences of untreated glaucoma, and the proper use of eye drops. The maximum possible score is 21, with a minimum of 0. The results of the GKQ-10 validation are presented in Table 1, with a Cronbach's $\alpha = 0.734$, indicating satisfactory reliability. The tool demonstrates adequate psychometric properties and is suitable for use among glaucoma patients. Sociodemographic and clinical data for the study population were collected through a custom questionnaire, hospital registry, and medical records.

The results of the survey analysis were subjected to statistical evaluation. The distribution of all quantitative parameters was tested for conformity with the normal distribution using the Kolmogorov–Smirnov and Shapiro–Wilk tests. For quantitative features, the following were calculated: the mean values (M), standard deviations (SD), medians (Me), lower (Q1) and upper quartiles (Q3), as well as extreme values: the minimum (Min) and maximum (Max). The significance of differences between average values for features with distributions that significantly deviated from normal or had heterogeneous variances between 2 groups was assessed using the Mann–Whitney U test. For nominal and ordinal qualitative variables, frequencies (n) and percentages (%) were calculated and presented in contingency tables (multidimensional). Hypotheses of no association between qualitative variables were tested using Pearson's χ^2 test or Fisher's exact test. A correlation between variables was considered significant if the test result was $p < 0.05$. Receiver operating characteristic (ROC) curves and Youden's index were used to determine the cutoff values for continuous and discrete variables. The diagnostic (classification) ability of the parameters that were analyzed was assessed using the area under the ROC curve (AUC). Additionally, sensitivity and specificity were estimated for dichotomous cutoff values. To assess the impact of quantitative variables on the likelihood of high adherence, the values of the univariate logistic regression coefficients (b) and their significance were estimated for each predictor. In all statistical tests, a significance level of $p < 0.05$ was considered statistically significant. The statistical analysis was conducted using Statistica v. 13.3 (TIBCO Software Inc., Tulsa, USA).

Results

Sociodemographic characteristics of the study population

A total of 190 individuals participated in the study, comprising 124 women (65.3%) and 66 men (34.7%), with an age range of 25–88 years (M = 69, SD = 11.2). The largest proportion of participants had completed secondary education (36.3%) or vocational training (32.1%), and the majority were retired at the time of data collection (77.9%). Most participants were in marital or domestic partnerships (65.8%), residing either with their partner (45.3%) or with other family members (44.2%). Additionally, 68.4% of the sample lived in urban areas. These sociodemographic data are summarized in Table 2.

Table 1. Validation results of the Glaucoma Knowledge Questionnaire (GKQ-10)

Item (question)	M*	SD*	r*	α*
1. Main source of knowledge about glaucoma (0–5 points)	10.0	2.2	0.283	0.512
2. Definition of the disease (0–3 points)	10.5	2.3	0.180	0.554
3. Possibility of a complete cure (0–1 point)	11.2	2.4	0.407	0.507
4. Possibility of total blindness (0–1 point)	11.2	2.5	0.356	0.519
5. Impact of medications on disease progression (0–1 point)	11.1	2.5	0.248	0.533
6. Possibility of discontinuing therapy (0–1 point)	11.1	2.5	0.392	0.526
7. Consequences of discontinuing therapy (0–1 point)	11.1	2.5	0.291	0.527
8. Importance of proper eye drop administration (0–1 point)	11.2	2.4	0.364	0.511
9. Time intervals between eye drop administrations (0–1 point)	11.3	2.4	0.409	0.493
10. Techniques for proper eye drop administration (0–6 points)	9.2	1.9	0.303	0.555

M = 12.0 points; SD = 2.6 points; Me [Q1; Q3] = 12 [10; 14]; n = 190; α = 0.734, average inter-item correlation r = 0.221, where: M* – mean after item removal, SD* – standard deviation after item removal, r* – item-scale linear correlation coefficient after item removal, α* – Cronbach's alpha after item removal. M – mean; Me – median.

Table 2. Sociodemographic characteristics of the study population

Characteristic (variable)	Total, n = 190	
	n	%
Gender		
Men	66	34.7
Women	124	65.3
Age [years]		
M (SD)	69.0 (11.2)	
Me [Q1; Q3]	70 [63; 77]	
Min–Max	25–88	
Education		
Primary	32	16.8
Vocational	61	32.1
Secondary	69	36.3
Higher	28	14.7
Marital status		
Single	65	34.2
In a relationship	125	65.8
Occupational activity		
Employed	31	16.3
Student	1	0.5
On disability pension	6	3.2
Retired	148	77.9
Unemployed	4	2.1
Place of residence		
Rural area	60	31.6
Urban area	130	68.4
Living status		
Living with spouse/partner	86	45.3
Living with family	84	44.2
Living alone	20	10.5

M – mean; SD – standard deviation; Me – median; Q1 – lower quartile; Q3 – upper quartile; Min – minimum value; Max – maximum value; n – number; % – percentage.

Clinical characteristics of the study population

The majority of participants had been diagnosed with glaucoma at least 10 years earlier (44.7%). All individuals were undergoing pharmacological treatment, with a subset also receiving adjunctive therapies, including laser treatment (10.5%) or surgical intervention (12.1%). Most of the patients were using a single anti-glaucoma medication (62.1%), typically administered twice daily (69.5%). Seventy-five percent of the patients had other systemic comorbidities. The detailed results are presented in Table 3,4.

Table 3. Clinical characteristics of glaucoma in the study population

Characteristic (variable)	Total, n = 190	
	n	%
Disease duration [years]		
M (SD)	11.3 (6.8)	–
Me [Q1; Q3]	10 [5; 15]	–
Min–Max	1–40	–
Disease duration		
1–5 years	48	25.3
6–10 years	57	30.0
More than 10 years	85	44.7
Type of glaucoma treatment*		
Pharmacological treatment	190	100.0
Laser therapy	20	10.5
Surgical intervention	23	12.1
Number of glaucoma medications currently in use		
1	118	62.1
2	62	32.6
3	10	5.3
Frequency of daily medication administration		
Once per day	51	26.8
Twice per day	132	69.5
Three times per day	7	3.7

Percentages do not add up to 100% as multiple responses were allowed. M – mean; SD – standard deviation; Me – median; Q1 – 1st quartile; Q3 – 3rd quartile; Min – minimum value; Max – maximum value.

Table 4. Clinical characteristics of patients – chronic diseases

Comorbidities	Total, n = 190	
	n	%
Hypertension	89	46.8
Diabetes	40	21.1
Asthma	5	2.6
Osteoporosis	5	2.6
Kidney failure	4	2.1
No other chronic conditions	47	24.7

Adherence to therapeutic recommendations among the glaucoma patients in the study

In the cohort of glaucoma patients examined, over half exhibited a low level of adherence to their prescribed treatment (58.9%). The mean adherence score, as assessed using the ARMS scale, was 16.3 ±2.7 points across the group (Table 5), highlighting a generally low adherence level within this population.

Table 5. Results of the assessment of adherence to pharmacological recommendations (Adherence to Refills and Medications Scale (ARMS)) in glaucoma patients

Variable	M (SD)	Me [Q1; Q3]	Min–Max
Total score	16.3 (2.7)	16 [14; 18]	12–25
Adherence level	n (%)		
Low (16–48 score)	112 (58.9%)		
High (12–15 score)	78 (41.1%)		

M – mean; SD – standard deviation; Me – median; Q1 – 1st quartile; Q3 – 3rd quartile; Min – minimum value; Max – maximum value.

Assessment of therapeutic communication between patients and physicians, measured with the CAT-14 questionnaire, in the glaucoma patient group

An analysis of communication quality indicated a high level of patient satisfaction with their interactions with the physician. The mean score across the 14 items of the questionnaire was 3.30 ±0.69. However, only 3.7% of responses rated the communication as excellent, giving the maximum possible score of 70 points (5 points per item). These findings are presented in Table 6.

The CAT-14 questionnaire results revealed that the highest percentage of excellent ratings was observed for the following: question 12, which addressed the physician's discussion of future steps and follow-up plans with the patient (10.0%); question 2, related to treating the patient with respect (9.5%); and question 7, concerning the provision of all the information that was expected (9.5%). The lowest percentage of excellent ratings was associated with questions about the physician's interest in the patient's opinion on their health (question 3), encouraging the patient to ask questions (question 10), and involving the patient in decision-making (question 11). The detailed results are presented in Table 7.

Table 6. Results of the assessment of patient–physician communication quality (Communication Assessment Tool (CAT-14))

Variable	M (SD)	Me [Q1; Q3]	Min–Max
CAT-14 (mean score of 14 items)	3.30 (0.69)	3.1 [3.0; 3.6]	1.7 – 5.0
Excellent ratings (CAT-14 = 5)	7 (3.7%)		

M – mean; SD – standard deviation; Me – median; Q1 – 1st quartile; Q3 – 3rd quartile; Min – minimum value; Max – maximum value.

Assessment of the patient–physician relationship, measured with the PDRQ-9 questionnaire, in the glaucoma patient group

In the studied group of glaucoma patients, the mean score for the patient–physician relationship, as measured using the PDRQ-9 questionnaire, was 30.4 ±6.2, indicating a moderate overall rating of the relationship. The analysis revealed that the largest proportion of respondents rated their relationship with their ophthalmologist as poor (42.1%). These results are summarized in Table 8.

The evaluation of the patient–physician relationship using the PDRQ-9 questionnaire allowed the entire cohort to be divided into 3 subgroups based on the PDRQ-9 scores:

– Group I – poor rating (PDRQ-9 sten scores 1–4), n = 80 patients;

– Group II – average rating (PDRQ-9 sten scores 5 and 6), n = 65 patients;

– Group III – good rating (PDRQ-9 sten scores 7–10), n = 45 patients.

Table 7. Results of the assessment of patient–physician communication quality ((Communication Assessment Tool (CAT-14))

Item (question)	M	SD	Me [Q1; Q3]	n	%
1. Greeted me in a way that made me feel comfortable	3.33	0.72	3 [3; 4]	15	7.9
2. Treated me with respect	3.38	0.72	3 [3; 4]	18	9.5
3. Showed interest in my opinion about my health	3.27	0.76	3 [3; 4]	13	6.8
4. Understood my main health concerns	3.31	0.77	3 [3; 4]	16	8.4
5. Paid attention to me (looked at me, listened carefully)	3.30	0.81	3 [3; 4]	16	8.4
6. Did not interrupt me while I was speaking	3.32	0.77	3 [3; 4]	17	8.9
7. Provided me with all the information I expected	3.41	0.82	3 [3; 4]	18	9.5
8. Spoke in a way I could understand	3.23	0.79	3 [3; 4]	16	8.4
9. Made sure I understood everything	3.29	0.79	3 [3; 4]	14	7.4
10. Encouraged me to ask questions	3.05	0.86	3 [3; 3]	13	6.8
11. Allowed me to participate in decision-making	3.13	0.86	3 [3; 4]	13	6.8
12. Discussed the next steps with me, including follow-up plans	3.46	0.81	3 [3; 4]	19	10.0
13. Showed care and concern for me	3.33	0.82	3 [3; 4]	14	7.4
14. Spent enough time with me	3.34	0.85	3 [3; 4]	16	8.4

M – mean; SD – standard deviation; Me – median; Q1 – lower quartile; Q3 – upper quartile; n – number of excellent ratings; % – percentage of excellent ratings.

Table 8. Results of the assessment of the patient–ophthalmologist relationship (PDQR-9)

Variable		M (SD)	Me [Q1;Q3]	Min–Max
Total score		30.4 (6.2)	28.5 [27; 33]	14–45
Quality of relationship (sten)		5.5 (2.0)	6 [4; 7]	1–10
Relationship rating, n (%)	Poor (1–4 sten)	80 (42.1)		
	Average (5, 6 sten)	65 (34.2)		
	Good (7–10 sten)	45 (23.7)		

M – mean; SD – standard deviation; Me – median; Q1 – lower quartile; Q3 – upper quartile; Min – minimum value; Max – maximum value.

Assessment of quality of life, measured with the NEI VFQ-25 questionnaire, in the glaucoma patient group

In the studied group of glaucoma patients, the mean score for the impact of visual impairment and symptoms on quality of life, as measured using the NEI VFQ-25 questionnaire, was 73.5 ±16.9 points. This finding indicates a relatively high quality of life, with limited impact of visual symptoms on patients' daily functioning. Only 2.1% of respondents reported a poor level of visual function, while 35.3% rated their visual function as average and 62.6% assessed it as good. The detailed results are summarized in Table 9.

The evaluation of visual function using the NEI VFQ-25 questionnaire allowed the cohort to be categorized into 3 subgroups based on NEI VFQ-25 scores:

– Group I – poor visual function (NEI VFQ 0–40 points), n = 4 patients;

– Group II – average visual function (NEI VFQ 40–70 points), n = 67 patients;

– Group III – good visual function (NEI VFQ 70–100 points), n = 119 patients.

Table 9. Results of visual function assessment National Eye Institute Visual Function Questionnaire (NEI VFQ-25)

Variable		M (SD)	Me [Q1; Q3]	Min–Max
Total score		73.5 (16.9)	77 [58; 87]	29.8–100
Visual function level, n (%)	Poor (0–40 points)	4 (2.1)		
	Average (40–70 points)	67 (35.3)		
	Good (70–100 points)	119 (62.6)		

M – mean; SD – standard deviation; Me – median; Q1 – lower quartile; Q3 – upper quartile; Min – minimum value; Max – maximum value.

Assessment of knowledge levels using the GKQ-10 questionnaire in the glaucoma patient group

The evaluation of glaucoma-related knowledge was performed using a custom-designed tool. The total knowledge score, with scores ranging from 0 to 21 points, was calculated based on responses to 10 questions. The raw scores were converted into sten scores (ranging from 1 to 10), and each participant was then classified into 1 of 3 categories based on their knowledge level: low, average, or high.

Table 10. Results of the glaucoma knowledge assessment (Glaucoma Knowledge Questionnaire (GKQ-10))

Variable		M (SD)	Me [Q1; Q3]	Min–Max
Glaucoma knowledge (total score)		12.0 (2.6)	12 [10; 14]	4.5–18
Glaucoma knowledge (sten scores)		5.5 (2.0)	6 [4; 7]	1–10
Knowledge level, n (%)	Low (1–4 sten)	59 (31.1%)		
	Average (5, 6 sten)	65 (34.2%)		
	High (7–10 sten)	66 (34.7%)		

M – mean; SD – standard deviation; Me – median; Q1 – lower quartile; Q3 – upper quartile; Min – minimum value; Max – maximum value.

The findings demonstrate that a significant proportion of patients had either a high (34.7%) or an average (34.2%) level of knowledge about their condition. A smaller proportion of patients had a low level of glaucoma-related knowledge (31.1%). The raw scores, sten scores, and distribution of patients across knowledge categories are detailed in Table 10.

The assessment of knowledge, as measured using the GKQ-10 questionnaire, allowed the study cohort to be divided into 3 subgroups based on GKQ-10 results:

– Group I – low knowledge level (GKQ-10 sten scores 1–4), n = 59 patients;

– Group II – average knowledge level (GKQ-10 sten scores 5 and 6), n = 65 patients;

– Group III – high knowledge level (GKQ-10 sten scores 7–10), n = 66 patients.

The impact of patient satisfaction with therapeutic communication on adherence to therapeutic recommendations in glaucoma patients

The association between the quality of communication between patients and their ophthalmologists, assessed using the CAT-14 questionnaire, and adherence to therapeutic recommendations, measured using the ARMS questionnaire, was statistically significant ($p < 0.01$). A weak but statistically significant negative correlation was observed between adherence to therapeutic recommendations (ARMS) and physician–patient communication (CAT), indicating that patients who rated communication with their physician more highly tended to adhere more closely to therapeutic guidelines.

Receiver operating characteristic curve analysis was used to determine the cutoff value for communication satisfaction (Fig. 1). For CAT scores ≥3.25, test sensitivity was 62.8%, specificity was 68.8%, and the AUC was 0.664. The lower limit of the 95% confidence interval (95% CI) for the AUC was 0.583, exceeding the acceptable threshold of 0.5, which indicates that a communication satisfaction–based test may be useful in classifying patients as adherent or non-adherent to therapeutic recommendations.

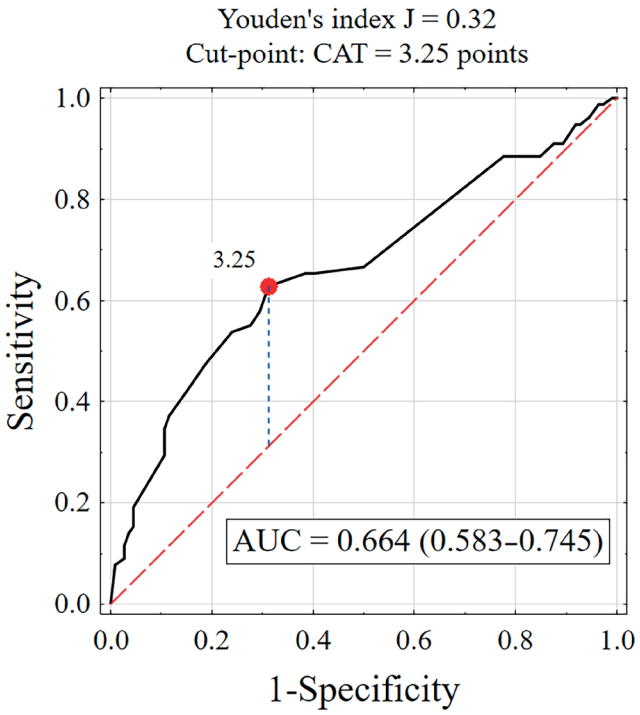

Fig. 1. Receiver operating characteristic (ROC) curve and communication cutoff value distinguishing between adherent and non-adherent patients. For Communication Assessment Tool (CAT-14) ≥ 3.25 points, the sensitivity of the communication-based test is 62.8% and specificity is 68.8%. Area under the ROC curve (AUC): 0.664

Comparative analysis revealed that patients with high adherence (ARMS ≤ 15) reported higher communication satisfaction than those with low adherence (ARMS > 15), with mean scores of 3.53 ±0.75 vs 3.13 ±0.59. Among those with high adherence, 31.2% expressed greater satisfaction with communication (CAT ≥ 3.25 points), while 37.2% of those with low adherence reported lower satisfaction (CAT < 3.25 points). These findings indicate that satisfaction with physician communication, as measured using the CAT questionnaire, is a statistically significant factor influencing adherence to therapeutic recommendations ($p < 0.001$), particularly with CAT scores ≥ 3.25 ($p < 0.001$). The detailed results are shown in Supplementary Table 1.

Univariate logistic regression analysis indicated that both higher physician–patient communication scores (CAT-14) and a CAT-14 score ≥3.25 were significant positive predictors of high adherence to therapeutic recommendations. In patients with CAT-14 scores ≥3.25, the likelihood of high adherence was nearly 4 times higher than in patients with lower communication scores (odds ratio (OR) = 3.72). These findings are presented in Supplementary Table 2.

The impact of the patient–physician relationship on adherence to therapeutic recommendations in glaucoma patients

The association between the quality of the patient–ophthalmologist relationship, as assessed using the PDRQ-9 questionnaire, and adherence to therapeutic recommendations, measured with the ARMS questionnaire, was not statistically significant when raw scores were analyzed ($p > 0.05$).

The influence of the patient–physician relationship on adherence to therapeutic recommendations appears to be minimal. The lower limit of the 95% CI for the ROC curve was below 0.5 (AUC = 0.476), indicating weak discriminatory ability for classifying patients as adherent or non-adherent based on this criterion. The results are presented in Fig. 2.

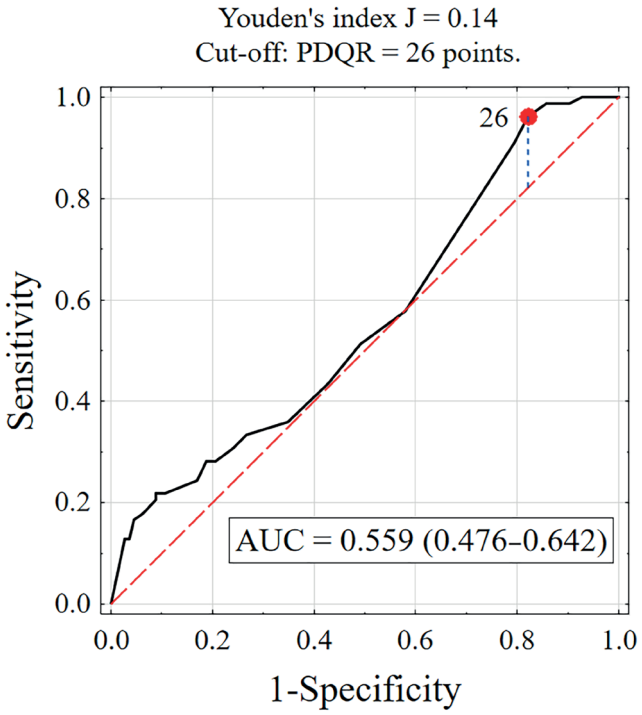

Fig. 2. Receiver operating characteristic (ROC) curve and cutoff point for the patient–physician relationship that differentiates between adherent and non-adherent patients. For Patient–Doctor Relationship Questionnaire (PDRQ- 9) ≥26 points, the test sensitivity for communication assessment is 96.2%, and specificity is 17.9%. Area under the ROC curve (AUC): 0.559

A comparative analysis showed that patients with high adherence (ARMS ≤ 15) reported better relationships with their physician compared to those with low adherence (ARMS > 15), with scores of 31.7 ±6.6 vs 29.5 ±5.8. Most patients with high adherence (96.2%) had a better patient–physician relationship (PDRQ-9 ≥ 26 points), while only 17.9% of those with low adherence had a poorer relationship with their physician (PDRQ-9 < 26 points). The analysis indicated that the patient–physician relationship, as measured by the PDRQ-9 questionnaire, is not a statistically significant predictor of adherence ($p = 0.163$). However, a relationship score of PDRQ-9 ≥ 26 points may be associated with adherence in this patient group ($p < 0.003$). The results are summarized in Supplementary Table 3.

Univariate logistic regression analysis showed that both a better patient–physician relationship, as measured using the PDRQ-9, and a PDRQ-9 score ≥ 26 are significant

positive predictors of adherence to therapeutic recommendations. Among patients with a PDRQ-9 score ≥26, the likelihood of high adherence was more than 5 times higher than among those with lower scores (OR = 5.44). However, the 95% CI (1.56–19.0) was quite wide.

The impact of quality of life on adherence to therapeutic recommendations in glaucoma patients

The association between visual function, as measured using the NEI VFQ-25 questionnaire, and adherence to therapeutic recommendations, measured with the ARMS questionnaire, was of limited significance for the raw data. The Spearman's rank correlation coefficient (rho = –0.069) did not differ significantly from zero.

Receiver operating characteristic curve analysis was used to determine the cutoff value for visual function (Fig. 3). For NEI VFQ scores ≥ 68, test sensitivity was 75.6% and specificity was 42.0%, with an AUC of 0.585. The lower limit of the 95% CI for the AUC was 0.502, which exceeded the accepted threshold of 0.5, suggesting that a visual function–based test may be useful in distinguishing between patients who adhere to therapeutic recommendations and those who do not.

Comparative analysis showed that patients with high adherence (ARMS ≤ 15) reported higher visual function scores than those with low adherence (ARMS > 15), with scores of 76.2 ±17.0 vs 71.7 ±16.7, respectively. The majority of patients with high adherence (71.8%) rated their visual function as good (NEI VFQ-25 ≥ 70 points), whereas less than half (42%) of patients with low adherence rated their visual function as average (NEI VFQ-25 40–70 points). Only 1.8% (n = 2) of patients with low adherence rated their visual function as poor (NEI VFQ-25 0–40 points). These findings indicate that visual function, as measured by the NEI VFQ-25, is a statistically significant factor influencing adherence to therapeutic recommendations (p = 0.048). Detailed results are summarized in Supplementary Table 5.

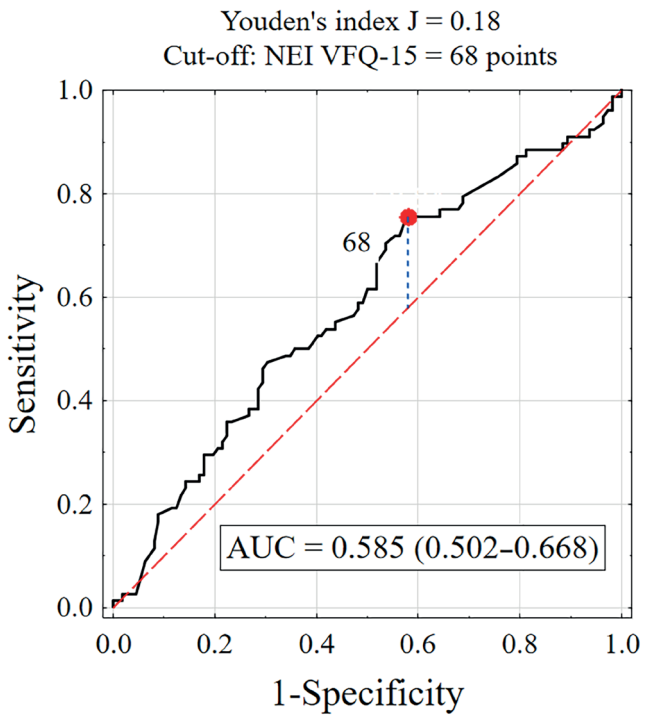

Fig. 3. Receiver operating characteristic (ROC) curve and threshold value for visual function (68 points and above), distinguishing between patients adhering and non-adhering to therapeutic recommendations. For National Eye Institute Visual Function Questionnaire (NEI VFQ-25) ≥68 points, the sensitivity of the visual function-based test is 75.6%, with a specificity of 42.0%. Area under the ROC curve (AUC): 0.585

Univariate logistic regression analysis showed that a higher visual function score (NEI VFQ-25 ≥ 68 points) is a significant positive predictor of adherence to therapeutic recommendations (b = 0.81; p = 0.013). The results are summarized in Supplementary Table 4.

The impact of glaucoma knowledge on adherence to therapeutic recommendations in glaucoma patients

The relationship between glaucoma knowledge, assessed using the custom GKQ-10 scale, and adherence to therapeutic recommendations, measured with the ARMS questionnaire, was statistically significant (p < 0.05). A weak but statistically significant negative correlation was observed between adherence to therapeutic recommendations (ARMS) and glaucoma knowledge (GKQ-10). Patients with greater glaucoma knowledge demonstrated higher adherence to therapeutic recommendations. Specifically, an increase of 1 point on the GKQ-10 glaucoma knowledge scale was associated with an average reduction of 0.27 points in the ARMS adherence score.

Receiver operating characteristic curve analysis was used to establish the cutoff value for glaucoma knowledge (Fig. 4). For GKQ-10 scores ≥ 13.5, test sensitivity was 48.7% and specificity was 75.0%, with an AUC of 0.631. The lower limit of the 95% CI for the AUC was 0.550, which exceeded the acceptable threshold of 0.5, indicating that a glaucoma knowledge–based test may be useful in classifying patients as adherent or non-adherent to therapeutic recommendations.

Comparative analysis showed that patients with higher adherence (ARMS ≤ 15) had greater glaucoma knowledge than those with lower adherence (ARMS > 15), with scores of 12.7 ±2.4 vs 11.5 ±2.6, respectively. Nearly half of patients with high adherence (48.7%) had a higher level of glaucoma knowledge (GKQ-10 ≥ 13.5), while the majority of patients with low adherence (75.0%) had lower levels of knowledge (GKQ-10 < 13.5). These findings indicate that glaucoma knowledge, as measured using the GKQ-10 questionnaire, is a statistically significant factor influencing adherence to therapeutic recommendations (p = 0.002). Furthermore, a GKQ-10 score ≥13.5 may be associated with adherence levels in this patient group (p = 0.002). The results are summarized in Supplementary Table 7.

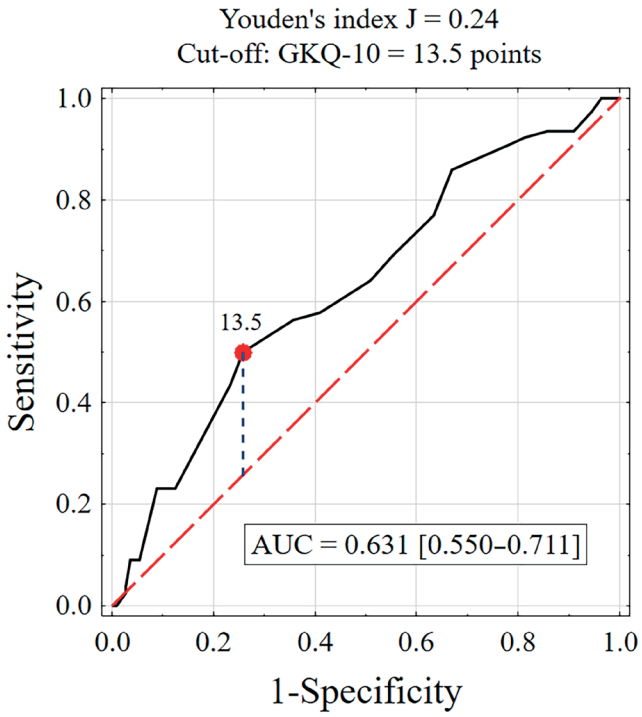

Fig. 4. Receiver operating characteristic (ROC) curve and threshold value for glaucoma knowledge (≥13.5 points) distinguishing between adherent and non-adherent patients. For Glaucoma Knowledge Questionnaire (GKQ-10) ≥13.5 points, the sensitivity of the knowledge-based test is 48.7%, and specificity is 75%. Area under the ROC curve (AUC): 0.631

Univariate logistic regression analysis showed that both higher levels of glaucoma knowledge, as measured using the GKQ-10 scale, and a cutoff value of GKQ-10 ≥ 13.5 are positive predictors of adherence to therapeutic recommendations. In patients with a GKQ-10 score of 13.5 or higher, the odds of high adherence were nearly 3 times higher than in those with lower knowledge (OR = 2.85). These findings are summarized in Supplementary Table 8.

Discussion

The study showed that more than half of the glaucoma patients (58.9%) exhibited low adherence to their prescribed therapy. This finding is consistent with previous reports in the literature. Systematic reviews indicate that medication non-adherence rates may reach up to 80%, although numerous studies report rates that are typically closer to 30%.[26,27] Schwartz and Quigley noted that while patients frequently self-report high adherence, objective measures such as pharmacy refill data or electronic monitoring often show much lower adherence levels. Adherence rates measured using electronic monitoring have been reported to be as low as 76%, depending on the method used.[27] Such discrepancies between self-reported adherence and actual behavior highlight the challenge of accurately assessing and addressing non-adherence in glaucoma patients. Factors such as the chronic and often asymptomatic nature of glaucoma contribute to this suboptimal adherence, as patients may not recognize the importance of consistent medication use.

The side effects associated with topical glaucoma medications, such as ocular irritation, hyperemia, and discomfort, may also represent significant barriers to adherence. In the study by Movahedinejad et al., adverse effects of glaucoma medications were identified as a significant factor contributing to non-adherence.[28] However, that study did not specify the exact nature of the side effects, but highlighted their significant role in adherence behaviors. Such adverse effects are particularly problematic for elderly patients, who may experience additional challenges with self-administering eye drops due to reduced manual dexterity or visual impairment. These issues may contribute to patients discontinuing or irregularly using their prescribed medication. In a cross-sectional survey, Newman-Casey et al. identified key barriers to adherence in glaucoma patients, including difficulties with eye drop administration and challenges related to medication scheduling. Difficulty administering eye drops was reported by 18% of participants who struggled with controlling the number of drops, 24% who had difficulty aiming, and others who experienced problems such as squeezing the bottle or holding it steady. Moreover, difficulties adhering to the prescribed medication schedule, often linked to forgetfulness, were associated with increased odds of non-adherence.[29]

The present study identified the quality of communication between patients and physicians as a key factor influencing adherence. Patients who reported higher satisfaction with communication, as measured with CAT-14 scores, demonstrated better adherence to their treatment regimen. Specifically, a CAT score ≥3.25 points was associated with a significantly higher likelihood of adherence. These findings underscore the importance of patient education about the disease and ensuring clear, empathetic communication between healthcare providers and patients.

Haskard Zolnierek and DiMatteo demonstrated in a meta-analysis that good physician–patient communication significantly enhances adherence, while poor communication increases the likelihood of non-adherence by 19%. Additionally, training physicians in communication skills can improve adherence by as much as 62%, underscoring that the quality of interaction is a modifiable factor that directly influences patient outcomes.[30] In a more specific clinical context, Sleath et al. examined patients with glaucoma and found that communication about how to properly administer eye drops was crucial to adherence. Their study showed that patients who were educated about medication administration and had higher self-efficacy were more likely to adhere to their prescribed regimen. Importantly, the study also highlighted racial disparities, with African American patients showing lower adherence, suggesting a need for communication strategies tailored to individual patient populations.[31] Carpenter

et al. expanded on these findings by investigating how communication affects glaucoma patients' self-efficacy, a key predictor of adherence. They found that provider education on glaucoma-related topics and discussions about patient views of the disease significantly improved patients' confidence in self-administering their medications. However, they also observed that patients who asked more medication-related questions tended to have lower confidence in overcoming barriers to adherence. This finding suggests that patients who are more engaged may simultaneously feel less assured about managing their treatment, indicating the need for deeper, problem-solving conversations during clinical visits.[32]

Although the quality of the patient–physician relationship, as measured using the PDRQ-9 questionnaire, did not show a statistically significant overall association with adherence, a better relationship (PDRQ-9 score ≥26) may still be positively associated with adherence. This finding warrants a nuanced interpretation. One possible explanation lies in the characteristics of the instrument itself. Although the PDRQ-9 has demonstrated strong psychometric properties, it may not fully capture specific aspects of the physician–patient relationship that are most relevant in chronic ophthalmic conditions, particularly those that are largely asymptomatic, such as glaucoma. Elements such as shared decision-making, patient activation, or trust in treatment recommendations may be underrepresented.

Alternatively, this finding may reflect the characteristics of the study population. In our cohort, many patients had been living with glaucoma for many years and may have developed established routines or views regarding their treatment independently of their current relationship with their physician. Furthermore, given the relatively high overall satisfaction with communication observed in the CAT-14 scores, the variability in PDRQ-9 responses may have been insufficient to detect a statistically significant effect. Future studies may benefit from using complementary instruments or qualitative approaches to further explore how relational factors influence adherence behaviors in this clinical context.

Gómez and Aillach examined various dimensions influencing the patient–physician relationship, highlighting the role of verbal and non-verbal communication skills, including clinical empathy, in improving patient outcomes. The review emphasized that while physician empathy and communication skills are associated with patient satisfaction and some clinical markers, further research is required to confirm these findings across a broader range of clinical outcomes.[33]

In their study, Stryker et al. found that non-adherent patients were less likely to be satisfied with the amount of time the physician spent discussing their eye condition, compared to adherent patients.[34] Patients emphasized the importance of physicians taking sufficient time to listen to their concerns and adequately address their questions. When patients felt that their doctors rushed through consultations or did not engage with their concerns, this created a barrier to adherence. A trusting relationship, built on empathy and respect, may help patients feel more confident in their treatment plan and more likely to adhere to prescribed regimens.[35]

The present study showed that patients with better visual function, as assessed using the NEI VFQ-25 questionnaire, were more likely to adhere to therapy. High adherence was associated with higher self-reported visual function (scores ≥68 points).

Jannuzzi et al. examined adherence to medication and its association with health-related quality of life (HRQoL) in elderly patients with diabetic retinopathy. In this cross-sectional study involving 100 elderly participants, the authors found that 58% adhered to at least 80% of their prescribed medication doses of antihypertensive and/or antidiabetic drugs. The Morisky Medication Adherence Scale indicated that one of the key factors contributing to non-adherence was discontinuation of medication after experiencing adverse drug effects. Additionally, the study also showed that patients with lower scores on the NEI VFQ-25, reflecting poorer vision-related quality of life, were more likely to be non-adherent.[36]

A statistically significant association was found between patients' knowledge about glaucoma, as measured using the GKQ-10 scale, and their adherence to therapy. Patients with higher levels of disease-related knowledge were more likely to adhere to therapeutic recommendations. A GKQ-10 score ≥13.5 points was associated with better adherence.

Patients' understanding and awareness of glaucoma are pivotal in shaping their adherence to therapy. Many glaucoma patients struggle to fully understand the progressive and asymptomatic nature of the disease, which often leads to underestimating its seriousness.[11] In their study, Birhanie et al. reported that patients with good knowledge about glaucoma were 2.24 times more likely to adhere to their treatment compared with patients with inadequate knowledge.[37]

In a study by Hoevenaars et al., the relationship between glaucoma-related knowledge and treatment adherence was investigated. Contrary to expectations, the study found no significant correlation between overall levels of glaucoma-related knowledge and adherence. Only specific knowledge items, such as understanding that eye drops cannot repair damage caused by glaucoma, were positively associated with adherence. This suggests that educational interventions focused solely on increasing knowledge may not substantially improve adherence in glaucoma patients.[10]

Welge-Lussen et al. conducted a study evaluating adherence among glaucoma patients to topical eye drops using self-reported questionnaires. The authors found no significant association between adherence and demographic factors, clinical characteristics, or glaucoma-related knowledge. The most common reason for non-adherence was forgetfulness. The findings suggest that individualized strategies are required to improve adherence, as general educational interventions may be insufficient in this context.[38]

Inadequate knowledge and reliance on a single source of information may result in poor adherence, as patients may not appreciate the importance of consistent, lifelong management of the disease.

A strong association has been reported between poor health literacy and non-adherence to glaucoma therapy. Many patients lack a clear understanding of their disease and the importance of adhering to treatment, particularly as glaucoma is often asymptomatic in its early stages. This lack of awareness may lead to missed doses and infrequent refills. Educational interventions tailored to patients' literacy levels – such as written instructions, educational videos, and personalized counseling – may significantly improve adherence. Patients who are educated about the proper administration of eye drops and the long-term consequences of non-adherence are more likely to adhere to their treatment regimens.[39,40]

In a study by Paczkowska et al., patient knowledge about arterial hypertension was significantly associated with treatment adherence and improved clinical outcomes, including better blood pressure control and reduced hospitalization rates. The authors found that patients with higher levels of education and those treated in specialized clinics had higher levels of knowledge and adherence to their prescribed therapy.[41]

Karaeren et al. examined the impact of patient knowledge on medication adherence in hypertensive individuals. The study found that specific aspects of knowledge, such as understanding the cause of hypertension and target blood pressure levels, were significantly associated with improved adherence. In contrast, knowledge of medication side effects was associated with decreased adherence.[42]

Figueira et al. examined the impact of educational interventions on disease-related knowledge, medication adherence, and glycemic control in patients with type 2 diabetes mellitus. The study showed significant improvements in knowledge ($p < 0.001$), adherence to medication ($p = 0.0318$), and glycemic control ($p = 0.0321$), suggesting that targeted educational strategies may effectively enhance patient self-management.[43]

Physical limitations, such as arthritis or other mobility issues, may make it difficult for patients to self-administer eye drops properly.[44] Additionally, complex treatment regimens that require multiple medications or frequent dosing can negatively affect adherence. Some patients report difficulty with eye drop bottle mechanisms or find it challenging to maintain a strict dosing schedule. However, the relationship between regimen complexity and adherence is not fully understood, as some studies have shown that patients receiving more complex regimens may adhere better, possibly because their disease is more advanced and because they receive increased counseling from healthcare providers.[13]

The findings of this study have important practical implications for ophthalmic care. Given the significant association between treatment adherence and both patient knowledge and the quality of physician–patient communication, there is a clear rationale for the implementation of targeted, low-cost interventions aimed at enhancing these aspects. Brief, structured educational sessions delivered during routine clinical visits may help patients better understand the chronic nature of glaucoma and the importance of consistent medication use, even in the absence of symptoms. Furthermore, communication skills training for ophthalmologists focused on empathetic engagement and effective information delivery may strengthen therapeutic relationships and positively influence adherence.

Such approaches are consistent with existing evidence from interventions in chronic disease management, including the Support, Educate, Empower (SEE) program developed by Newman-Casey et al., which has demonstrated improvements in adherence through personalized coaching and motivational interviewing.[45] Similarly, a systematic review by Nieuwlaat et al. underscores that tailored, theory-informed interventions delivered by trained health professionals can be modestly effective in improving medication adherence.[46] Future implementation of such strategies in glaucoma care may help mitigate disease progression and reduce the societal burden of vision loss.

Limitations of the study

Several limitations should be considered when interpreting the results of this study. First, the cross-sectional design of the study does not allow determination of causal relationships between the examined variables and adherence to treatment. Second, the study was conducted in a single ophthalmology clinic in Poland, which may limit the generalizability of the findings to other populations or healthcare settings. Additionally, patient self-reporting of treatment adherence and the use of questionnaires to assess communication, relationship quality, and knowledge may introduce bias, such as social desirability or recall bias. Further longitudinal studies are necessary to confirm these findings and explore the long-term impact of the identified factors on adherence to glaucoma treatment.

Conclusions

Improving adherence to glaucoma medications requires addressing a range of factors, including demographic disparities, financial barriers, patient education, health beliefs, and the physical demands of treatment. Tailored interventions that focus on patient education, cost reduction strategies, and simplified treatment regimens are essential for enhancing adherence and improving long-term outcomes for patients with glaucoma.

This study demonstrated that adherence to glaucoma treatment is influenced by a range of factors, including patient knowledge about the disease, satisfaction with

doctor–patient communication, and visual functioning. Patients who reported better communication and higher levels of knowledge about glaucoma were significantly more likely to adhere to their treatment regimens. Although the quality of the patient–physician relationship did not show a statistically significant effect on adherence, a positive relationship may still contribute to better adherence outcomes.

These findings highlight the need for targeted interventions that focus on improving patient education, addressing misconceptions about glaucoma, and fostering empathetic, clear communication between patients and healthcare providers. Educational efforts should emphasize the importance of lifelong treatment for glaucoma, even in the absence of symptoms, to prevent disease progression and irreversible vision loss.

Supplementary files

The supplementary materials are available at https://doi.org/10.5281/zenodo.15552922. The package contains the following files:

Supplementary Table 1. Assessment of patient–physician communication satisfaction (CAT-14) in groups with different levels of adherence to therapeutic recommendations.

Supplementary Table 2. Logistic regression results for adherence to therapeutic recommendations (ARMS ≤ 15 points) in patient groups with different levels of communication with ophthalmologists (CAT) and ORs with 95% CIs.

Supplementary Table 3. Assessment of the patient–physician relationship (PDRQ-9) in groups with different levels of adherence to therapeutic recommendations.

Supplementary Table 4. Logistic regression results for adherence to therapeutic recommendations (ARMS ≤ 15 points) in patient groups with different levels of communication with ophthalmologists (PDRQ-9) and ORs with 95% CIs.

Supplementary Table 5. Assessment of visual function in groups with different levels of adherence to therapeutic recommendations.

Supplementary Table 6. Logistic regression analysis of adherence to therapeutic recommendations (ARMS ≤ 15 points) in patient groups differing in visual function, including ORs and 95% CIs.

Supplementary Table 7. Assessment of glaucoma knowledge level (GKQ-10) in groups with different levels of adherence to therapeutic recommendations.

Supplementary Table 8. Logistic regression results for adherence to therapeutic recommendations (ARMS ≤ 15 points) in patient groups with differing levels of glaucoma knowledge, including ORs and 95% CIs.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Zenodo repository at https://doi.org/10.5281/zenodo.15552500.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Katarzyna Malewicz https://orcid.org/0000-0002-8609-8285
Mariusz Chabowski https://orcid.org/0000-0002-9232-4525
Jakub Staś https://orcid.org/0000-0002-5072-1290
Anna Maria Cybulska https://orcid.org/0000-0002-6912-287X
Anna Szymańska-Chabowska https://orcid.org/0000-0003-3615-6923
Beata Jankowska-Polańska https://orcid.org/0000-0003-1120-3535

References

1. Stein JD, Khawaja AP, Weizer JS. Glaucoma in adults-screening, diagnosis, and management: A review. *JAMA*. 2021;325(2):164–174. doi:10.1001/jama.2020.21899
2. GBD 2019 Blindness and Vision Impairment Collaborators, Vision Loss Expert Group of the Global Burden of Disease Study. Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020. The Right to Sight: An analysis for the Global Burden of Disease Study. *Lancet Glob Health*. 2021;9(2):e144–e160. doi:10.1016/S2214-109X(20)30489-7
3. McMonnies CW. Glaucoma history and risk factors. *J Optom*. 2017;10(2):71–78. doi:10.1016/j.optom.2016.02.003
4. Soh Z, Yu M, Betzler BK, et al. The global extent of undetected glaucoma in adults. *Ophthalmology*. 2021;128(10):1393–1404. doi:10.1016/j.ophtha.2021.04.009
5. Anton A, Serrano D, Nolivos K, et al. Cost-effectiveness of screening for open angle glaucoma compared with opportunistic case finding. *J Glaucoma*. 2023;32(2):72–79. doi:10.1097/ijg.0000000000002132
6. Sultan MB, Mansberger SL, Lee PP. Understanding the importance of IOP variables in glaucoma: A systematic review. *Surv Ophthalmol*. 2009;54(6):643–662. doi:10.1016/j.survophthal.2009.05.001
7. Heijl A, Leske MC, Bengtsson B, et al. Reduction of intraocular pressure and glaucoma progression: Results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol*. 2002;120(10):1268–1279. doi:10.1001/archopht.120.10.1268
8. Kass MA, Heuer DK, Higginbotham EJ, et al. The Ocular Hypertension Treatment Study: A randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. *Arch Ophthalmol*. 2002;120(6):701–713; discussion 829–830. doi:10.1001/archopht.120.6.701
9. Vrijens B, De Geest S, Hughes DA, et al. A new taxonomy for describing and defining adherence to medications. *Br J Clin Pharmacol*. 2012;73(5):691–705. doi:10.1111/j.1365-2125.2012.04167.x
10. Hoevenaars JGMM, Schouten JSAG, Van Den Borne B, Beckers HJM, Webers CAB. Will improvement of knowledge lead to improvement of compliance with glaucoma medication? *Acta Ophthalmol*. 2008;86(8):849–855. doi:10.1111/j.1755-3768.2007.01161.x
11. Newman-Casey PA, Shtein RM, Coleman AL, Herndon L, Lee PP. Why patients with glaucoma lose vision: The patient perspective. *J Glaucoma*. 2016;25(7):e668–675. doi:10.1097/IJG.0000000000000320
12. Zaharia AC, Dumitrescu OM, Radu M, Rogoz RE. Adherence to therapy in glaucoma treatment: A review. *J Pers Med*. 2022;12(4):514. doi:10.3390/jpm12040514
13. Moore SG, Richter G, Modjtahedi BS. Factors affecting glaucoma medication adherence and interventions to improve adherence: A narrative review. *Ophthalmol Ther*. 2023;12(6):2863–2880. doi:10.1007/s40123-023-00797-8
14. Uchino M, Yokoi N, Shimazaki J, Hori Y, Tsubota K; on behalf of the Japan Dry Eye Society. Adherence to eye drops usage in dry eye patients and reasons for non-compliance: A web-based survey. *J Clin Med*. 2022;11(2):367. doi:10.3390/jcm11020367

15. Oltramari L, Mansberger SL, Souza JMP, De Souza LB, De Azevedo SFM, Abe RY. The association between glaucoma treatment adherence with disease progression and loss to follow-up. *Sci Rep*. 2024;14(1):2195. doi:10.1038/s41598-024-52800-2
16. Alghamdi A, Keegan D, Connell P, Dooley I, O'Toole L. Adherence of patients with age-related macular degeneration to AREDS 2-recommended nutritional supplements. *Ir J Med Sci*. 2023;192(6):3163–3167. doi:10.1007/s11845-023-03351-4
17. Kripalani S, Risser J, Gatti ME, Jacobson TA. Development and evaluation of the Adherence to Refills and Medications Scale (ARMS) among low-literacy patients with chronic disease. *Value Health*. 2009; 12(1):118–123. doi:10.1111/j.1524-4733.2008.00400.x
18. Lomper K, Chabowski M, Chudiak A, Białoszewski A, Dudek K, Jankowska-Polańska B. Psychometric evaluation of the Polish version of the Adherence to Refills and Medications Scale (ARMS) in adults with hypertension. *Patient Prefer Adherence*. 2018;12:2661–2670. doi:10.2147/ppa.s185305
19. Makoul G, Krupat E, Chang CH. Measuring patient views of physician communication skills: Development and testing of the Communication Assessment Tool. *Patient Educ Couns*. 2007;67(3):333–342. doi:10.1016/j.pec.2007.05.005
20. Świątoniowska-Lonc N, Białoszewski A, Makoul G, Jankowska-Polańska B. Translation and cultural adaptation of the Polish version of the Communication Assessment Tool (CAT). *Risk Manag Healthc Policy*. 2020;13:1533–1542. doi:10.2147/rmhp.s261710
21. Van der Feltz-Cornelis CM, Van Oppen P, Van Marwijk HWJ, De Beurs E, Van Dyck R. A patient–doctor relationship questionnaire (PDRQ-9) in primary care: Development and psychometric evaluation. *Gen Hosp Psychiatry*. 2004;26(2):115–120. doi:10.1016/j.genhosppsych.2003.08.010
22. Porcerelli JH, Murdoch W, Morris P, Fowler S. The Patient–Doctor Relationship Questionnaire (PDRQ-9) in primary care: A validity study. *J Clin Psychol Med Settings*. 2014;21(3):291–296. doi:10.1007/s10880-014-9407-2
23. Zenger M, Schaefert R, Van Der Feltz-Cornelis C, Brähler E, Häuser W. Validation of the Patient–Doctor-Relationship Questionnaire (PDRQ-9) in a representative cross-sectional German population survey. *PLoS One*. 2014;9(3):e91964. doi:10.1371/journal.pone.0091964
24. Brola W, Opara J, Fudala M, Czernicki J, Szwejkowski W. Polish adaptation and validation of the Visual Function Questionnaire (VFQ-25) in multiple sclerosis patients [in Polish]. *Wiad Lek*. 2010;63(3):161–170. PMID:21125739.
25. Revicki DA, Rentz AM, Harnam N, Thomas VS, Lanzetta P. Reliability and validity of the National Eye Institute Visual Function Questionnaire-25 in patients with age-related macular degeneration. *Invest Ophthalmol Vis Sci*. 2010;51(2):712. doi:10.1167/iovs.09-3766
26. Olthoff C, Schouten J, Vandeborne B, Webers C. Noncompliance with ocular hypotensive treatment in patients with glaucoma or ocular hypertension: An evidence-based review. *Ophthalmology*. 2005;112(6):953–961.e7. doi:10.1016/j.ophtha.2004.12.035
27. Schwartz GF, Quigley HA. Adherence and persistence with glaucoma therapy. *Surv Ophthalmol*. 2008;53(Suppl 1):S57–S68. doi:10.1016/j.survophthal.2008.08.002
28. Movahedinejad T, Adib-Hajbaghery M. Adherence to treatment in patients with open-angle glaucoma and its related factors. *Electron Physician*. 2016;8(9):2954–2961. doi:10.19082/2954
29. Newman-Casey PA, Robin AL, Blachley T, et al. The most common barriers to glaucoma medication adherence: A cross-sectional survey. *Ophthalmology*. 2015;122(7):1308–1316. doi:10.1016/j.ophtha.2015.03.026
30. Haskard Zolnierek KB, DiMatteo MR. Physician communication and patient adherence to treatment: A meta-analysis. *Med Care*. 2009; 47(8):826–834. doi:10.1097/mlr.0b013e31819a5acc
31. Sleath B, Blalock SJ, Carpenter DM, et al. Ophthalmologist–patient communication, self-efficacy, and glaucoma medication adherence. *Ophthalmology*. 2015;122(4):748–754. doi:10.1016/j.ophtha.2014.11.001
32. Carpenter DM, Blalock SJ, Sayner R, et al. Communication predicts medication self-efficacy in glaucoma patients. *Optom Vis Sci*. 2016; 93(7):731–737. doi:10.1097/opx.0000000000000856
33. Gómez G, Aillach E. Ways to improve the patient–physician relationship. *Curr Opin Psychiatry*. 2013;26(5):453–457. doi:10.1097/yco.0b013e328363be50
34. Stryker JE, Beck AD, Primo SA, et al. An exploratory study of factors influencing glaucoma treatment adherence. *J Glaucoma*. 2010; 19(1):66–72. doi:10.1097/ijg.0b013e31819c4679
35. Piette JD, Heisler M, Krein S, Kerr EA. The role of patient–physician trust in moderating medication nonadherence due to cost pressures. *Arch Intern Med*. 2005;165(15):1749. doi:10.1001/archinte.165.15.1749
36. Jannuzzi FF, Cintra FA, Rodrigues RCM, São-João TM, Gallani MCBJ. Medication adherence and quality of life among the elderly with diabetic retinopathy. *Rev Lat Am Enfermagem*. 2014;22(6):902–910. doi:10.1590/0104-1169.3477.2494
37. Birhanie SA, Getie GA, Tesfa M, et al. Treatment adherence and associated factors among glaucoma patients attending ophthalmic units of referral hospitals in North West Ethiopia, 2019. *Front Ophthalmol (Lausanne)*. 2022;2:985893. doi:10.3389/fopht.2022.985893
38. Yu A, Welge-Lussen U, Weise S. Assessing the adherence behavior of glaucoma patients to topical eye drops. *Patient Prefer Adherence*. 2014;9:17–23. doi:10.2147/ppa.s69943
39. Juzych MS, Randhawa S, Shukairy A, Kaushal P, Gupta A, Shalauta N. Functional health literacy in patients with glaucoma in urban settings. *Arch Ophthalmol*. 2008;126(5):718–724. doi:10.1001/archopht.126.5.718
40. Muir KW, Santiago-Turla C, Stinnett SS, et al. Health literacy and adherence to glaucoma therapy. *Am J Ophthalmol*. 2006;142(2):223–226.e2. doi:10.1016/j.ajo.2006.03.018
41. Paczkowska A, Hoffmann K, Kus K, et al. Impact of patient knowledge on hypertension treatment adherence and efficacy: A single-centre study in Poland. *Int J Med Sci*. 2021;18(3):852–860. doi:10.7150/ijms.48139
42. Karaeren H, Yokuşoğlu M, Uzun S, et al. The effect of the content of the knowledge on adherence to medication in hypertensive patients. *Anadolu Kardiyol Derg*. 2009;9(3):183–188. PMID:19520651.
43. Figueira ALG, Boas LCGV, Coelho ACM, Freitas MCFD, Pace AE. Educational interventions for knowledge on the disease, treatment adherence and control of diabetes mellitus. *Rev Lat Am Enfermagem*. 2017;25:e2863. doi:10.1590/1518-8345.1648.2863
44. Adamson E, Kendall G. Difficulty in eye drop administration for people with rheumatoid arthritis. *Br J Occup Ther*. 2016;79(9):550–556. doi:10.1177/0308022616643101
45. Newman-Casey PA, Resnicow K, Winter S, et al. The Support, Educate, Empower personalized glaucoma coaching trial design. *Clin Trials*. 2023;20(2):192–200. doi:10.1177/17407745221136571
46. Nieuwlaat R, Wilczynski N, Navarro T, et al. Interventions for enhancing medication adherence. *Cochrane Database Syst Rev*. 2014;2014(11): CD000011. doi:10.1002/14651858.cd000011.pub4

Electromyographic activity of the orbicularis oris muscle in children with and without lip competence: A cross-sectional study

Liliana Szyszka-Sommerfeld[1,A–F], Monika E. Machoy[2,C,F], Jacek Świtała[3,F], Magdalena Sycińska-Dziarnowska[3,F], Krzysztof Woźniak[3,E,F], Gianrico Spagnuolo[4,5,E,F], *Luigi Esposito[4,E,F], *Carlo Rengo[4,E,F]

[1] Laboratory for Propaedeutic of Orthodontics and Facial Congenital Defects, Pomeranian Medical University in Szczecin, Poland
[2] Department of Periodontology, Pomeranian Medical University in Szczecin, Poland
[3] Department of Maxillofacial Orthopedics and Orthodontics, Pomeranian Medical University in Szczecin, Poland
[4] Department of Neurosciences, Reproductive and Odontostomatological Sciences, University of Naples "Federico II", Italy
[5] School of Dentistry, College of Dental Medicine, Kaohsiung Medical University, Taiwan

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(5):809–818

Address for correspondence
Liliana Szyszka-Sommerfeld
E-mail: liliana.szyszka.sommerfeld@pum.edu.pl

Funding sources
None declared

Conflict of interest
None declared

Acknowledgements
The first author would like to express sincere gratitude to Gianrico Spagnuolo for his supervision and guidance in the preparation of this article. This work was developed within the framework of a scientific research collaboration during a Scientific Internship completed by the first author at the Department of Neurosciences, Reproductive and Odontostomatological Sciences, University of Naples "Federico II", Italy.

*Luigi Esposito and Carlo Rengo contributed equally to this work.

Received on February 20, 2025
Reviewed on June 30, 2025
Accepted on July 18, 2025
Published online on April 14, 2026

Cite as
Szyszka-Sommerfeld L,Machoy ME, Świtała J, et al. Electromyographic activity of the orbicularis oris muscle in children with and without lip competence: A cross-sectional study. *Adv Clin Exp Med*. 2026;35(5):809–818. doi:10.17219/acem/208381

DOI
10.17219/acem/208381

Abstract

Background. The "equilibrium theory" posits that the tongue and perioral muscles, including the orbicularis oris (OO) muscle, function synergistically to maintain balanced tooth positioning. Surface electromyography (sEMG) is a valuable, nonivasive method for assessing muscle activity. However, previous electromyographic (EMG) studies comparing lip muscle activity in children with and without lip competence (LC) have yielded inconsistent results. Therefore, further research is needed to clarify OO muscle activity patterns in this population.

Objectives. The aim of the study was to evaluate the EMG activity of the superior (SOO) and inferior orbicularis oris (IOO) muscles in children with and without LC.

Materials and methods. The sample comprised 30 children with lip incompetence (LI) (mean age 9.46 ± 1.76 years) and 30 children with LC (mean age 8.85 ± 1.52 years). Electromyographic recordings of the SOO and IOO muscles were obtained using a DAB Bluetooth Instrument (Zebris Medical GmbH, Isny im Allgäu, Germany) at clinical rest, during saliva swallowing, lip protrusion ("kissing" position), lip compression, and while articulating the syllables /pa/, /ba/, and /ma/. Statistical analyses were performed using Stata v. 11.0 (StataCorp, College Station, USA). The level of significance was set at $p < 0.05$.

Results. Electromyographic activity of the SOO and IOO muscles during saliva swallowing ($p < 0.001$, adjusted p (p_{adj}) = 0.002) and lip compression ($p = 0.001$, $p_{adj} = 0.013$ for SOO; $p < 0.001$, $p_{adj} = 0.005$ for IOO) was significantly greater in children with LI compared to those with LC. Similar EMG activity at rest and during speech production was observed in children with and without LC.

Conclusions. Children with LI demonstrate increased SOO and IOO muscle activity during saliva swallowing and lip compression, suggesting greater muscular effort is required to achieve lip seal. This increased activity may disturb the muscular force balance essential for proper maxillofacial growth and could contribute to the development of malocclusion.

Key words: surface electromyography, craniofacial growth, lip competence, lip incompetence, incompetent lips

Highlights

- Children with lip incompetence exhibit increased orbicularis oris muscle activity during saliva swallowing and lip compression.
- This increased activity may indicate a need for additional effort to achieve lip closure and could influence maxillofacial morphology.
- The findings underscore the importance of early diagnosis and intervention for lip competence, as functional lip alterations may occur even in the absence of overt malocclusions.

Background

Competent lips are essential for maintaining a balance between the muscles of the cheeks and the tongue. Lip competence (LC) is defined as the ability to maintain slight contact between the lips in a state of clinical rest (relaxed musculature). In contrast, the term incompetent lips (IL) describes anatomically short lips that are unable to maintain contact in a state of relaxed musculature. In such cases, lip seal requires active contraction of the orbicularis oris (OO) and mentalis muscles, clinically observed as chin wrinkling.[1–5] Incompetent lips may result from anteroposterior and vertical disharmony of the dentofacial complex or may develop as an adaptation to mouth breathing due to nasal obstruction, respiratory difficulties, or other forms of altered oral function.[5]

The "equilibrium theory" proposes that the tongue and perioral muscles, including the OO, work together to maintain balanced tooth positioning.[6,7] Proper growth and development of the maxillofacial structures and stable occlusion depend on a balance of muscular forces. Maintaining healthy perioral soft tissue function is essential, as lip and tongue dysfunction may lead to malocclusion or orthodontic relapse.[8–13]

Surface electromyography (sEMG) is a valuable tool for assessing muscle activity. This noninvasive technique measures electrical signals on the skin above superficial muscles, providing an objective, quantitative representation of neuromuscular balance in the stomatognathic system. Therefore, it is a highly useful instrument for studying the relationship between morphology and function in the oral-maxillofacial system.[14–19] The sEMG's simplicity, nonivasiveness, and accessibility are important factors for studies involving children.[20–23]

Currently, there is limited research on OO muscle EMG activity in patients with lip incompetence (LI). Moreover, existing studies vary with respect to study populations, methodologies (e.g., tasks assessed using sEMG), and results. Some studies have focused on EMG activity of the OO in children and adolescents.[1,2,10,24–27] Some authors have reported EMG findings in patients with LI and malocclusion,[1,24–27] while others did not assess occlusal characteristics in their study populations.[2,10] The results of previous studies vary considerably.

Some studies reported higher EMG activity of the OO during tasks such as swallowing,[1,2,24–26] lip compression,[24–27] speech,[2] chewing,[2] puffing out the cheeks,[10] and at rest[5] in individuals with LI. Other investigations reported lower EMG activity of the inferior OO at rest and during swallowing in individuals with LI, or similar activity during speech and swallowing in the superior OO in children with and without IL.[10] Because EMG studies of lip activity comparing children with CL and IL have reported contradictory results, and EMG patterns of the OO, particularly in children with LI, remain unclear, further research on lip EMG activity in this population is necessary.

Objectives

The objective of this study was to evaluate the EMG activity of the superior (SOO) and inferior (IOO) orbicularis oris muscles in children with and without IL. We hypothesized that there would be no difference in the EMG activity of the SOO and IOO muscles between children with and without LC during the studied tasks.

Materials and methods

This cross-sectional study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement. The study adhered to the Declaration of Helsinki and protocol was approved by the Local Bioethics Committee of the Pomeranian Medical University in Szczecin, Poland (approval No. KB-0012/08/15). Written informed consent was obtained from the parents before the procedures.

Study population

The study included 30 children with IL (mean age: 9.46 ±1.76 years) and 30 children with CL (mean age: 8.85 ±1.52 years). The children were recruited over 2 months from a consecutive sample at an orthodontic clinic. They were selected in the order of arrival for screening, forming a convenience sample. During the clinical examination, each child sat upright without head support,

looking straight ahead with the jaw relaxed. An orthodontist classified each child into either the competent lip group (CLG), defined as lips lightly touching without mentalis muscle contraction, or the incompetent lip group (ILG), defined as lips apart at rest or lips touching with visible mentalis muscle activity (chin shrinkage).[2]

The study included both boys and girls with mixed dentition, Class I occlusion, and positive overjet and overbite. Children had to be free of congenital disorders, behavioral problems, disabilities, environmental allergies, common colds, nasal obstruction, oral parafunctional habits (e.g., thumb or finger sucking, nail biting), restrictions of the labial or lingual frenulum (e.g., ankyloglossia), and atypical swallowing (AS). They must not have undergone previous orthodontic treatment, myofunctional or speech therapy, or trauma or surgery in the orofacial region. Children with systemic diseases affecting muscles, muscle disorders such as temporomandibular dysfunction or bruxism, or a history of neuromuscular disease were excluded. Participants must not have been taking any medications that might affect muscle activity. Only children for whom voluntary parental consent was obtained were included. Children who did not meet these criteria were excluded from the study.

Sample size

Each study group was planned to include 30 participants. The sample size was determined using a two-sided t-test with a significance level of 0.05 and a power of 80% ($1 - \beta = 0.20$). The effect size, defined as large (Cohen's $d = 0.80$), was determined based on pilot data and prior research.[2,10,25] The calculation was performed using the pwr v. 1.3.0 and report v. 0.5.7 packages in R v. 4.3.1 (R Foundation for Statistical Computing, Vienna, Austria), running on Windows 10 x64 (Microsoft Corp., Redmond, USA).[28,29] An additional 15% correction was applied to account for deviations from a normal distribution.

Experimental procedures

An experienced orthodontist determined whether each participant had CL or IL during the clinical examination. The examination also included an assessment of breathing mode (via a parental questionnaire), mouth posture (observed by the examiner), ear, nose and throat conditions, and speech and language. Perioral muscles were visually assessed to diagnose AS, defined as the tongue pressing against the dental arches or being positioned anteriorly or laterally during swallowing and accompanied by myofunctional alterations, such as hyperactivation of the lips, facial muscle tension, or abnormal head and mandible movements. An intraoral analysis evaluated occlusal features, including vertical overlap, overjet, Angle Class, crossbite, and open bite. Body mass index (BMI) was calculated for each participant using weight [kg] and height [m].

To assess the correlation between LC and dentofacial morphology, cephalometric radiographs routinely obtained during the diagnostic process were analyzed. All cephalometric radiographs were acquired using a digital X-ray device (Cranex Tom; Soredex, Tuusula, Finland). The measurements were evaluated by the same orthodontist using cephalometric analysis software (Orthodontics 7.0; Orto-Bajt, Wrocław, Poland) and were based on the Segner and Hasund method.[30]

Surface electromyography

The EMG activity of the SOO and IOO muscles was simultaneously measured using a DAB Bluetooth electromyography device (Zebris Medical GmbH, Isny im Allgäu, Germany) with a gain of 1000, a high-pass filter of 7–5 kHz, a sampling rate of 1 kHz, and a 12-bit analog-to-digital converter. The system was calibrated before each recording. Bipolar Ag/AgCl surface electrodes (Noraxon Dual Electrode; Noraxon, Scottsdale, USA) with a 20-mm inter-electrode distance were placed on the SOO and IOO muscles. Electromyographic procedures followed established protocols.[24–26] During the recording, participants sat in a dental chair with their heads in a natural position. Electrodes were positioned as follows: SOO muscle – along the line from the lip corner to the subnasal point; IOO muscle – along the line from the lip corner to the chin midline; reference electrode – behind the right ear. A single trained and experienced examiner, blinded to the clinical findings, positioned the electrodes to ensure consistent placement relative to muscle fiber orientation.

To prepare for EMG recordings, each participant's skin was gently cleaned with a 70% ethyl alcohol solution and dried using disposable cotton to reduce impedance. Slight skin reddening following cleaning was considered an indicator of adequate preparation. After electrode placement, an impedance test was conducted using a Metex P-10 measuring device (Metex Instruments Corporation, Seoul, South Korea). Surface EMG recordings of the SOO and IOO muscles commenced 5 min later.

The EMG recordings of the SOO and IOO muscles were obtained while participants performed a series of tasks: 1) resting with relaxed lips; 2) swallowing saliva; 3) pursing the lips into a "kissing" position; 4) compressing the lips together; and 5) pronouncing the syllables /pa/, /ba/, and /ma/. Before the EMG recordings, the examiner explained each task to ensure proper execution. Participants were provided with instructions and asked to practice the tasks by imitating the examiner. They were instructed to: 1) relax the jaw; 2) swallow saliva naturally (as visually confirmed by hyoid bone movement); 3) pucker the lips into a "kissing" position; 4) press the lips together firmly; and 5) articulate the syllables /pa/, /ba/, and /ma/.

Following established protocols, each movement was repeated at least 3 times.[24–26] For each participant, the final EMG value represented the average of the last

2 recordings, as the initial measurement often deviated substantially and was discarded as a "learning" trial.[31] To prevent muscle fatigue, participants rested for approx. 1 min between tasks. Electrode signals were amplified, digitized, and stored on a computer. Surface EMG data were subsequently normalized to the peak EMG value for each task. Among various normalization protocols, some authors recommend using the peak EMG value obtained during dynamic contraction as the reference value, as this approach helps reduce inter-subject variability.[32]

The task with the highest potential served as the maximum reference point (taken as 100% OO muscle activity). Lip protrusion (pursing lips into a "kissing" position) showed significantly higher EMG activity than all other tasks and was used as the reference for normalization. Measured values were expressed as a percentage of the lip protrusion reference value using the formula: mean task activity (μV)/mean lip protrusion activity (μV) × 100%. This resulted in normalized EMG values [μV/μV%] for all measurements.[24–26,33,34]

The repeatability of the EMG protocol was assessed by performing duplicate measurements in 20 randomly selected children. The same examiner repeated the recordings after a 15-min interval under identical experimental conditions. The comparison of the 1st and 2nd measurements demonstrated good intra-examiner repeatability.

Statistical analyses

All statistical analyses were performed using Stata v. 11.0 (StataCorp, College Station, USA). The level of statistical significance was set at $p < 0.05$. To assess differences in EMG activity of the SOO and IOO muscles between the ILG and CLG across 6 functional tasks (rest, lip compression, swallowing saliva, lip protrusion, and articulation of /pa/, /ba/, and /ma/), a total of 12 primary comparisons were conducted (6 per muscle). To reduce the risk of type I error due to multiple testing, the Holm–Bonferroni correction was applied, maintaining the family-wise error rate at $\alpha = 0.05$.

This method ranks the unadjusted p-values from smallest (rank 1) to largest (rank 12) and compares each to a sequentially adjusted threshold according to the following equation:

$$\text{P-value (unadjusted)} \leq \text{Threshold}_k \quad (1),$$

where

$\text{Threshold}_k = \alpha/(m - k + 1)$ (1);

m – total number of comparisons (12);

k – rank of the unadjusted p-value (1 to 12).

An unadjusted p-value is considered statistically significant if it is less than or equal to its corresponding threshold. Because only 2 groups were compared, omnibus tests were not required. The Holm–Bonferroni correction was therefore applied directly to the 12 predefined pairwise comparisons constituting the primary endpoints. Exploratory subgroup analyses (by sex) were performed without p-value adjustment to preserve statistical power for hypothesis generation. However, these secondary analyses should be interpreted with caution due to the increased risk of type I error.[35]

Normality of distribution was assessed separately for each group using the Shapiro–Wilk test. For variables demonstrating normal distribution in both groups, homogeneity of variances was evaluated using Levene's test (Supplementary Tables 1–3 present the tests of distributional assumptions for variables included in Tables 1–3, respectively). Continuous variables meeting both normality and homogeneity of variance assumptions were analyzed using the independent-samples Student's t-test. In cases of unequal variances, Welch's t-test was applied. When normality was violated in at least 1 group, continuous variables were analyzed using the Mann–Whitney U test.

Table 1. Characteristics of the study population

Characteristics		ILG (n = 30)	CLG (n = 30)	Statistics	p-value
Age [years]		9.48 ±1.77	8.87 ±1.52	t(58) = 1.43	0.157#
Sex n (%)	boys	15 (50.0)	13 (43.3)	$\chi^2 = 0.27$, df = 1	0.605∂
	girls	15 (50.0)	17 (56.7)		
BMI [kg/m^2]		21.20 (18.77–22.77)	22.35 (20.45–24.10)	U = 373, Z = –1.14	0.258*
Overjet [mm]		3.71 ±1.16	2.59 ±0.87	t(58) = 4.26	<0.001#
Overbite [mm]		1.55 (1.20–1.88)	1.70 (1.42–2.00)	U = 352.0	0.148*
ANB [°]		3.70 (3.50–3.88)	2.49 (2.45–2.52)	U = 868.5	<0.001*
Sp'–Gn [mm]		68.04 ±2.75	57.84 ±2.33	t(58) = 15.52	<0.001#
Index [%]		73.06 ±1.40	81.05 ±2.55	Welch t(47.8) = –15.02	<0.001&

Normally distributed continuous variables are presented as mean ± standard deviation (SD). Non-normally distributed variables are presented as median and interquartile range (IQR; Q1–Q3). Categorical variables are presented as counts and percentage (%). Statistics indicates the test value (t, Welch t, U, χ^2) and the number of degrees of freedom (df); #Student's t-test; &Welch's t-test; *Mann–Whitney U test; ∂ Pearson's χ^2 test.
ILG – incompetent lip group; CLG – competent lip group; BMI – body mass index; ANB – angle determined by the N-A and N-B lines (N – Nasion, the most anterior point on the frontonasal suture; A – the deepest point on the contour of the maxilla; B – the deepest point on the contour of the mandible); Sp'–Gn – distance between the Spina' and Gnathion points, lower anterior facial height (Sp'– point of intersection of lines Nasion–Gnathion and Spina nasalis anterior–Pterygomaxillare; Gnathion – inferior point on the mandibular symphysis); Index (ratio between the middle and lower facial heights N–Sp'/Sp'–Gn × 100%).

Table 2. Electrical activity of the SOO muscle [μV/μV%] in the ILG and CLG

Activity	Sex	ILG (n = 30)	CLG (n = 30)	Statistics	p	p_{adj}	Effect size
Rest	all	19.10 ±5.20	20.50 ±4.80	t(58) = 0.52	0.606#	>0.999	d = 0.13
	girls	18.80 ±4.90	20.40 ±4.70	t(28) = 0.31	0.760#	–	d = 0.11
	boys	23.90 ±5.50	20.30 ±4.90	t(28) = 0.65	0.520#	–	d = 0.24
Compression of the lips	all	133.20 ±12.50	99.00 ±14.10	t(58) = 3.35	0.001#	0.013	d = 0.86
	girls	122.20 ±10.20	105.90 ±12.30	t(28) = 0.15	0.880#	–	d = 0.05
	boys	143.30 ±14.40	92.20 ±13.60	t(28) = 1.55	0.130#	–	d = 0.57
Swallowing of saliva	all	98.44 (69.27–118.99)	55.49 (42.46–86.34)	U = 700.0	< 0.001*	0.002	r = –0.56
	girls	92.34 (62.06–115.74)	55.63 (41.39–82.32)	U = 110.0	0.590*	–	r = –0.12
	boys	100.99 (80.89–127.46)	55.35 (48.32–89.88)	U = 95.0	0.420*	–	r = –0.28
Articulating /pa/	all	101.97 (64.82–149.38)	76.90 (65.11–86.63)	U = 506.0	0.706*	>0.999	r = –0.06
	girls	99.56 (64.87–149.38)	81.58 (59.56–96.31)	U = 115.0	0.360*	–	r = –0.18
	boys	105.29 (57.78–153.85)	91.71 (66.36–158.21)	U = 98.0	0.850*	–	r = –0.04
Articulating /ma/	all	81.36 (56.93–137.14)	88.18 (69.24–113.74)	U = 459.0	0.900*	>0.999	r = –0.02
	girls	75.84 (52.45–128.93)	89.35 (70.60–114.87)	U = 108.0	0.480*	–	r = –0.14
	boys	101.97 (66.96–165.49)	87.99 (68.79–89.99)	U = 102.0	0.360*	–	r = –0.20
Articulating /ba/	all	97.77 (81.55–120.80)	93.35 (75.53–113.87)	U = 500.0	0.464*	>0.999	r = –0.08
	girls	102.40 (69.32–113.76)	93.23 (70.59–109.22)	U = 116.0	0.610*	–	r = –0.12
	boys	93.89 (84.15–149.07)	93.47 (79.25–119.65)	U = 105.0	0.750*	–	r = –0.06

Normally distributed continuous variables are presented as mean ± standard deviation (SD). Non-normally distributed variables are presented as medians (interquartile range (IQR); 1st and 3rd quartiles (Q1–Q3)). Statistics indicates the test value and the number of degrees of freedom (t(df), U); p – raw p-value (two-tailed Mann–Whitney U or Student's t tests); #Student's t-test; *Mann–Whitney U test; p_{adj} – Holm–Bonferroni-adjusted p-value (only presented for "all" rows); r – rank-biserial correlation (95% confidence interval (95% CI)); d – Cohen's d (95% CI); ILG – incompetent lips group; CLG – competent lips group. Rows labelled "all" are the 12 a priori EMG endpoints; their p_{adj} correct the family-wise error (FWER) at α = 0.05. The "girls" and "boys" rows are exploratory subgroup analyses; their p-values remain unadjusted. Descriptive effect sizes for "girls" and "boys" are also reported, but not subjected to the multiplicity correction. Effect sizes are given as rank-biserial r (nonparametric) or Cohen's d (parametric).

Table 3. Electrical activity of the IOO muscle [μV/μV%] in the ILG and CLG

Activity	Sex	ILG (n = 30)	CLG (n = 30)	Statistics	p-value	p_{adj}	Effect size
Rest	all	19.75 ±5.10	20.90 ±4.60	t(58) = 0.42	0.670#	>0.999	d = 0.10
	girls	19.70 ±4.20	20.80 ±4.10	t(28) = 0.02	0.989#	–	d = 0.01
	boys	23.70 ±5.80	20.98 ±4.70	t(28) = 1.27	0.205#	–	d = 0.30
Compression of the lips	all	133.45 ±15.20	99.95 ±18.40	t(58) = 3.50	< 0.001#	0.005	d = 0.90
	girls	130.60 ±14.50	105.40 ±16.20	t(28) = 0.27	0.791#	–	d = 0.07
	boys	143.50 ±17.80	94.50 ±15.30	t(28) = 1.58	0.122#	–	d = 0.60
Swallowing of saliva	all	98.20 (78.60–114.81)	63.80 (55.45–76.49)	U = 720.0	<0.001*	0.002	r = –0.55
	girls	94.70 (69.80–104.50)	67.80 (58.70–76.50)	U = 112.0	0.704*	–	r = –0.13
	boys	107.60 (78.60–125.40)	56.70 (54.50–74.30)	U = 98.0	0.299*	–	r = –0.30
Articulating /pa/	all	101.95 (67.90–149.70)	86.25 (65.40–116.50)	U = 510.0	0.412*	>0.999	r = –0.05
	girls	99.60 (65.45–149.70)	81.30 (59.70–96.50)	U = 118.0	0.296*	–	r = –0.19
	boys	106.50 (69.70–153.40)	91.20 (65.50–156.70)	U = 100.0	0.643*	–	r = –0.02
Articulating /ma/	all	85.95 (65.70–140.68)	88.40 (77.33–109.53)	U = 460.0	0.882*	>0.999	r = –0.01
	girls	75.60 (56.70–113.20)	90.50 (76.80–114.50)	U = 110.0	0.384*	–	r = –0.15
	boys	102.40 (65.70–165.40)	82.30 (78.90–89.70)	U = 104.0	0.215*	–	r = –0.21
Articulating /ba/	all	98.70 (72.30–149.80)	93.85 (73.40–116.40)	U = 510.0	0.442*	>0.999	r = –0.07
	girls	102.30 (70.90–113.50)	94.30 (70.50–109.30)	U = 120.0	0.558*	–	r = –0.11
	boys	87.90 (83.40–154.80)	93.40 (74.50–118.70)	U = 102.0	0.401*	–	r = –0.05

Normally distributed continuous variables are presented as mean ± standard deviation (SD). Non-normally distributed variables are presented as medians (interquartile range (IQR); 1st and 3rd quartiles (Q1–Q3)). Statistics indicates the test value and the number of degrees of freedom (t(df), U); p – raw p-value (two-tailed Mann–Whitney U or Student's t tests); #Student's t-test; *Mann–Whitney U test; p_{adj} – Holm–Bonferroni-adjusted p-value (only presented for "all" rows); r – rank-biserial correlation (95% confidence interval (95% CI)); d – Cohen's d (95% CI); ILG – incompetent lips group; CLG – competent lips group. Rows labelled "all" are the 12 a priori EMG endpoints; their p_{adj} correct the family-wise error (FWER) at α = 0.05. The "girls" and "boys" rows are exploratory subgroup analyses; their p-values remain unadjusted. Descriptive effect sizes for "girls" and "boys" are also reported, but not subjected to the multiplicity correction. Effect sizes are given as rank-biserial r (nonparametric) or Cohen's d (parametric).

Categorical variables were compared using Pearson's χ^2 test (Supplementary Table 4 presents the observed and expected frequencies, along with the results of Pearson's χ^2 test for the association between gender and LC). The paired Student's t-test was used to analyze the repeatability of the EMG procedure, as the derived data were normally distributed. Continuous variables were reported as mean ± standard deviation (SD) for normally distributed data or as median with interquartile range (IQR, 25^{th}–75^{th} percentiles) for non-normally distributed data. Categorical variables were presented as counts and percentages (%).

Results

A total of 60 individuals participated in the study. The ILG comprised 30 children (15 girls (50%) and 15 boys (50%) mean age: 9.46 ±1.76 years). The CLG included 30 subjects (17 girls (56.7%) and 13 boys (43.3%), mean age: 8.85 ±1.52 years). Table 1 presents the characteristics of the participants. The groups showed statistically significant differences in average overjet, ANB angle, Sp'–Gn distance, and Index ($p < 0.001$); however, no significant differences were found for the other analyzed variables.

The 1^{st} and 2^{nd} EMG examinations revealed no statistically significant differences in SOO and IOO activity. Children with IL exhibited significantly higher SOO and IOO muscle activity during swallowing compared to those with CL ($p < 0.001$, Holm–Bonferroni-adjusted $p_{adj} = 0.002$) (Tables 2,3; Fig. 1). Similarly, EMG activity of both the SOO and IOO muscles was significantly higher in the ILG during lip compression ($p = 0.001$, $p_{adj} = 0.013$ for SOO and $p < 0.001$, $p_{adj} = 0.005$ for IOO) (Tables 2,3; Fig. 2).

Resting SOO and IOO activity levels were similar in both groups ($p = 0.606$, $p_{adj} > 0.999$ for SOO and $p = 0.670$, $p_{adj} > 0.999$ for IOO) (Tables 2,3; Fig. 3). No statistically significant differences were found in SOO and IOO muscle activity during the production of bilabial sounds (/pa/, /ba/, /ma/) between the ILG and CLG (/pa/ $p = 0.706$, /ba/ $p = 0.464$, /ma/ $p = 0.900$, $p_{adj} > 0.999$ for SOO and /pa/ $p = 0.412$, /ba/ $p = 0.442$, /ma/ $p = 0.882$, $p_{adj} > 0.999$ for IOO) (Tables 2,3; Fig. 4–6). No statistically significant differences in SOO and IOO muscle activity during any of the tasks were observed between boys and girls in either group (Tables 2,3).

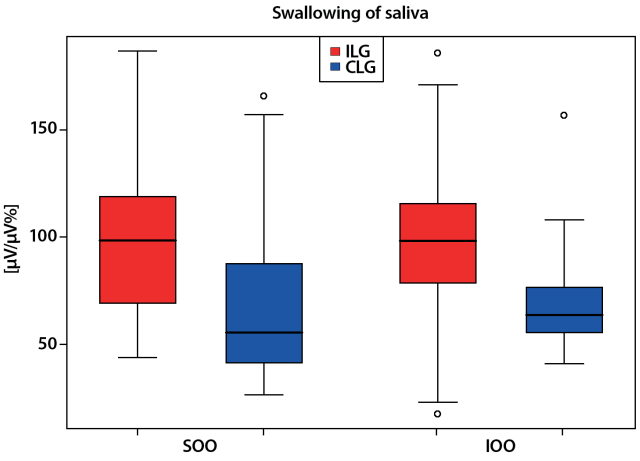

Fig. 1. Electromyographical (EMG) activity of the superior (SOO) and inferior (IOO) orbicularis oris muscles [µV/µV%] during saliva swallowing in the incompetent lips (ILG) and the competent lips (CLG) groups; $p < 0.001$; adjusted p-value (p_{adj}) = 0.002 for the SOO and $p < 0.001$; $p_{adj} = 0.002$ for the IOO muscle activity; box – interquartile range (IQR); horizontal line in the box – median; whiskers – min–max; dots – outliers

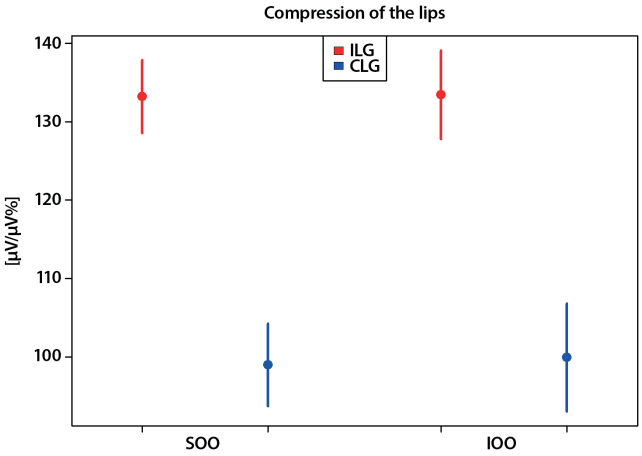

Fig. 2. Electromyographical (EMG) activity of the superior (SOO) and inferior (IOO) orbicularis oris muscles [µV/µV%] during lip compression in the incompetent lips (ILG) and the competent lips (CLG); $p = 0.001$; adjusted p-value (p_{adj}) = 0.013 for the SOO muscle activity and $p < 0.001$; $p_{adj} = 0.005$ for the IOO muscle activity; dot – mean; whiskers – 95% confidence interval (95% CI)

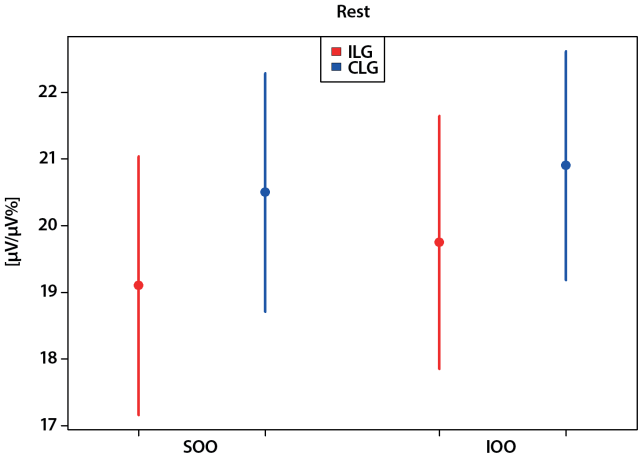

Fig. 3. Electromyographical (EMG) activity of the superior (SOO) and inferior (IOO) orbicularis oris muscles [µV/µV%] at rest in the incompetent lips (ILG) and the competent lips (CLG) groups; $p = 0.606$, adjusted p-value (p_{adj}) > 0.999 for the SOO muscle activity and $p = 0.670$; $p_{adj} > 0.999$ for the IOO muscle activity; dot – mean; whiskers – 95% confidence interval (95% CI)

Discussion

This study compared the electrical activity of the OO muscle in children with IL to that of children with CL

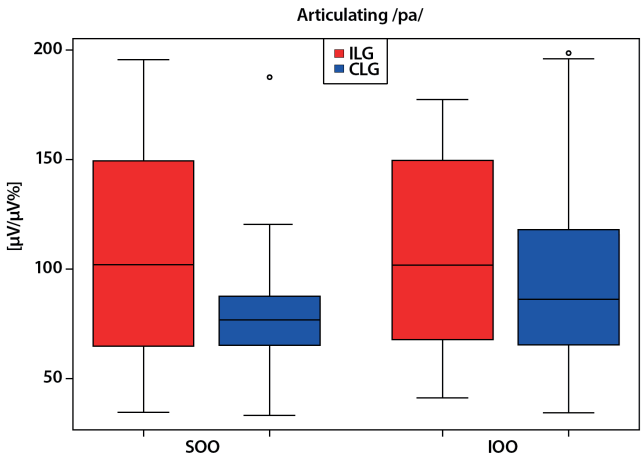

Fig. 4. Electromyographical (EMG) activity of the superior (SOO) and inferior (IOO) orbicularis oris muscles [µV/µV%] during articulating /pa/ in the incompetent lips (ILG) and the competent lips (CLG) groups; p = 0.706; adjusted p-value (p_{adj}) > 0.999 for the SOO muscle activity and p = 0.412; p_{adj} > 0.999 for the IOO muscle activity; box – interquartile range (IQR); horizontal line in the box – median; whiskers – min–max; dots – outliers

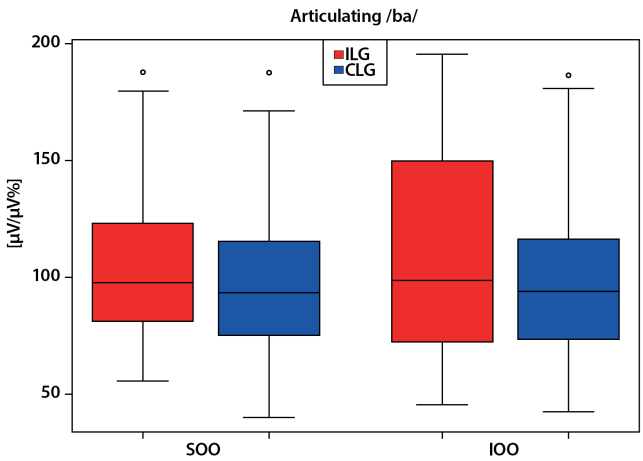

Fig. 5. Electromyographical (EMG) activity of the superior (SOO) and inferior (IOO) orbicularis oris muscles [µV/µV%] during articulating /ba/ in the incompetent lips (ILG) and the competent lips (CLG) groups; p = 0.464; adjusted p-value (p_{adj}) > 0.999 for the SOO muscle activity and p = 0.442; p_{adj} > 0.999 for the IOO muscle activity; box – interquartile range (IQR); horizontal line in the box – median; whiskers – min–max; dots – outliers

using surface electrodes. Repetition of the procedure confirmed the reliability of electrode placement and the study protocol. Higher activity of the SOO and IOO muscles was observed in children with IL during swallowing and lip compression, suggesting increased effort required to achieve lip seal. The significantly higher activity during swallowing may be particularly important due to the frequency of this action (600–3,000 times per day), potentially affecting maxillofacial morphology and occlusal stability. Previous research has demonstrated increased EMG activity of the OO muscle in individuals with LI compared to those with normal lip seal.[1,2,24–27] However, the findings remain inconsistent, likely due to heterogeneity in study

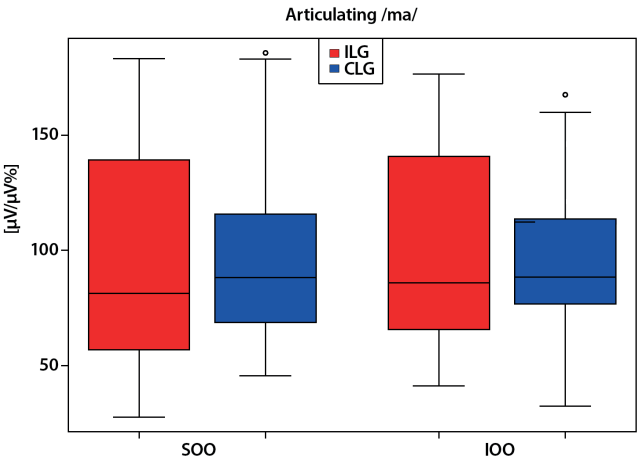

Fig. 6. Electromyographical (EMG) activity of the superior (SOO) and inferior (IOO) orbicularis oris muscles [µV/µV%] during articulating /ma/ in the incompetent lips (ILG) and the competent lips (CLG) groups; p = 0.900; adjusted p-value (p_{adj}) > 0.999 for the SOO muscle activity and p = 0.882; p_{adj} > 0.999 for the IOO muscle activity; box – interquartile range (IQR), horizontal line in the box – median; whiskers – min–max; dots – outliers

populations and methodological differences. Our results, showing higher OO muscle activity during swallowing in children with LI, are consistent with the observations of Gustafsson and Ahlgren[1] and Gamboa et al.,[2] although the latter study included an older cohort (17–27 years of age).

In contrast, Lipari et al.[10] reported similar SOO muscle activity in both groups and lower IOO activity during swallowing in children with IL aged 7–13 years. They proposed that children with IL have hypotonic IOO muscles and require activation of the mentalis muscle to achieve lip seal.[10] Consistent with our findings, Tosello et al.[27] observed significantly higher activity in the SOO, IOO, and mentalis muscles during lip compression in children with IL, suggesting muscle hyperactivity.

In the present study, similar OO muscle activity at rest was observed in children with and without IL, consistent with the findings of Gamboa et al.[2] However, this result contrasts with the findings of Tomiyama et al.,[5] who reported higher resting EMG activity in adults with IL. Lipari et al.[10] observed lower resting activity in children with IL. Orbicularis oris muscle activity during speech was similar in both groups, aligning with the results of Lipari et al.[10] This suggests that speech function in the OO muscle is not affected by LC, potentially due to compensatory mentalis muscle activity.[36] However, this differs from Gamboa et al.,[2] who found higher SOO activity during speech in individuals with IL.

When interpreting the study results, it is important to consider that various factors may influence the findings. Previous research has suggested that EMG activity in the SOO muscle is elevated in individuals with LI, as well as in those with anterior open bite and increased overjet.[37] In the present study, no study participants with open bite were included, and no significant difference

in mean overbite was observed between the groups. Although overjet was significantly greater in the ILG than in the CLG, no significant differences in OO muscle EMG activity at rest were observed between the groups, suggesting that overjet does not significantly affect this parameter.

Moreover, it should be noted that the study sample was controlled for malocclusion. Only participants with Class I occlusion and positive overjet and overbite were included. However, significant differences in cephalometric characteristics were observed, including a larger ANB angle, increased Sp'–Gn distance, and a reduced Index (indicating greater lower anterior facial height) in the ILG. These findings suggest that, independent of occlusal conditions, dentofacial morphology may be associated with LI, which is consistent with previous studies.[38] Another factor that could affect OO muscle activity is AS. Some authors[39,40] have reported higher EMG activity of the OO muscle in patients with AS and LI compared to subjects with LC and controls, suggesting greater effort during swallowing.

In our study, however, patients with clearly identifiable AS were excluded. Nevertheless, a few borderline cases were observed in the ILG, as these children exhibited some features of an infantile swallowing pattern. Many other factors, such as disabilities, oral parafunctional habits, restrictions of the labial or lingual frenulum, as well as a history of frenotomy or speech and myofunctional therapy, may also affect muscle activity.[25,41–45] It has been reported that higher resting activity of the perioral muscles is observed among thumb and pacifier suckers,[41] as well as in individuals with disabilities (e.g., Down syndrome and orally educated deaf individuals), during swallowing or speech.[25,42] Similarly, the presence of ankyloglossia may affect the resting position of the tongue and lips.[43] In light of the above, children with these characteristics were excluded from the study sample. It should also be noted that individuals with nasal obstruction or other airway-related conditions were excluded. However, some of these conditions may have been underreported, as the assessment was based on clinical examination and a parent-completed questionnaire.

The present study utilized sEMG to measure muscle activity. This method is particularly suitable for use in children, as it does not involve needles or other invasive procedures. However, sEMG recordings may be influenced by factors such as inconsistent electrical conductivity and difficulty in distinguishing signals from different muscle fibers.[15] Crosstalk, defined as the recording of signals over a muscle that originate from adjacent muscles, represents a significant source of error in sEMG and may lead to overestimation of muscle activity. Several factors influence the extent of crosstalk, including the inter-electrode distance of the EMG sensor and the accuracy of sensor placement on the muscle surface.[46,47]

The accuracy of sEMG recordings also depends on factors related to the psychological and physiological state of the participants, as well as the equipment used.[14,15] To address these challenges, a consistent inter-electrode distance was maintained, electrodes were carefully placed at the center of the muscle belly, and the data were analyzed using appropriate quantitative methods, including a normalization procedure, to enhance accuracy. Another limitation of sEMG is its sensitivity to impedance imbalance.

To minimize impedance, the skin was thoroughly cleaned with 70% ethyl alcohol, a method commonly employed in sEMG studies involving children.[24–26,45] The children included in this study were well accustomed to the procedure, and the use of alcohol did not cause discomfort or elicit abnormal muscle activity.

In summary, it is important for clinicians to recognize that increased muscle activity in patients with LI may eventually influence dental arch form and facial morphology. Furthermore, stable tooth positioning relies on a balance among the oral, maxillofacial, and respiratory systems. Therefore, the therapeutic goal should be to establish normal jaw posture, nasal breathing, and proper lip closure.[2,48,49]

The diagnosis of LC is a critical element of orthodontic treatment planning. In this context, sEMG may serve as a valuable noninvasive tool to support accurate diagnosis. A precise analysis of sEMG recordings of the OO muscles is essential for refining treatment protocols aimed at achieving orofacial muscular balance. In this regard, the present study highlights the importance of early diagnosis of LC and timely intervention strategies, as functional lip alterations may occur even in the absence of overt malocclusion.

If left untreated, these functional disturbances may lead to more complex orthodontic, periodontal, or esthetic complications over time. Therefore, knowledge of OO muscle activity in growing individuals with LI may be particularly valuable for clinicians involved in the multidisciplinary management of such cases.[10,25,50–53]

For growing patients diagnosed with LI, the combined use of myofunctional therapy and interceptive orthodontics – alongside conventional orthodontic treatment – is fundamental to establishing new neuromuscular patterns and restoring harmony within the stomatognathic system.[45,49,54,55] Surface electromyography may serve as a valuable tool for monitoring the progression and therapeutic outcomes of these interventions.[45]

Limitations of the study

This study has several limitations that should be acknowledged. We recognize that regulation of OO muscle activity is multifactorial and may be differentially influenced by occlusal conditions, oral habits, soft tissue morphology, and vertical and sagittal facial dimensions.[13,56] One limitation of the present study is the lack of detailed analysis of soft tissue morphology, such as upper lip length. In addition, the diagnosis of LI was based on clinical observation – specifically, the presence of lip separation at clinical rest or lip contact achieved through visible contraction of the mentalis muscle in the mandibular rest position (chin shrinkage). It should be noted that subtle contraction of the orofacial muscles may

go unnoticed during visual assessment.[4] Another limitation of this study is the lack of subdivision of the ILG into more specific subgroups, which could have allowed for a more detailed characterization of functional differences within this population. Furthermore, the cross-sectional design provides data collected at a single point in time and does not allow for conclusions regarding causality or longitudinal changes. Therefore, considering the aforementioned limitations, further well-designed, long-term studies conducted on larger sample sizes are required to confirm these findings and to better understand their potential impact on maxillofacial morphology.

Conclusions

Despite its limitations, the findings of this study indicate that children with LI exhibit increased EMG activity of the SOO and IOO muscles during saliva swallowing and lip compression. This may reflect the need for greater muscular effort to achieve adequate lip seal. Such increased muscular activity could potentially disturb the equilibrium of perioral forces required for proper maxillofacial growth and development, thereby contributing to the risk of malocclusion. However, longitudinal studies are necessary to confirm these observations and to clarify their long-term clinical implications.

Supplementary files

The supplementary files are available at https://doi.org/10.5281/zenodo.15807698. The package contains the following files:

Supplementary Table 1. Tests of distributional assumptions for variables in Table 1.

Supplementary Table 2. Tests of distributional assumptions for variables in Table 2.

Supplementary Table 3. Tests of distributional assumptions for variables in Table 3.

Supplementary Table 4. Observed and expected frequencies with Pearson's χ^2 test results for gender and LC association.

Data Availability Statement

The datasets supporting the findings of the current study are openly available in Zenodo repository at https://doi.org/10.5281/zenodo.16093588.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Liliana Szyszka-Sommerfeld https://orcid.org/0000-0002-1103-1297
Monika E. Machoy https://orcid.org/0000-0001-5787-222X
Jacek Świtała https://orcid.org/0009-0006-6468-9742
Magdalena Sycińska-Dziarnowska https://orcid.org/0000-0002-4794-8216
Krzysztof Woźniak https://orcid.org/0000-0002-5088-8760
Gianrico Spagnuolo https://orcid.org/0000-0003-3769-9786
Luigi Esposito https://orcid.org/0009-0003-1558-0423
Carlo Rengo https://orcid.org/0000-0003-1127-0877

References

1. Gustafsson M, Ahlgren J. Mentalis and orbicularis oris activity in children with incompetent lips: An electromyographic and cephalometric study. *Acta Odontol Scand*. 1975;33(6):355–363. doi:10.3109/00016357509004640
2. Gamboa NA, Miralles R, Valenzuela S, et al. Comparison of muscle activity between subjects with or without lip competence: Electromyographic activity of lips, supra- and infrahyoid muscles. *Cranio*. 2017;35(6):385–391. doi:10.1080/08869634.2016.1261441
3. Tosello DO, Vitti M, Berzin F. EMG activity of the orbicularis oris and mentalis muscles in children with malocclusion, incompetent lips and atypical swallowing: Part I. *J Oral Rehabil*. 1998;25(11):838–846. doi:10.1046/j.1365-2842.1998.00322.x
4. Yamaguchi K, Morimoto Y, Nanda RS, Ghosh J, Tanne K. Morphological differences in individuals with lip competence and incompetence based on electromyographic diagnosis. *J Oral Rehabil*. 2000;27(10):893–901. doi:10.1111/j.1365-2842.2000.00596.x
5. Tomiyama N, Ichida T, Yamaguchi K. Electromyographic activity of lower lip muscles when chewing with the lips in contact and apart. *Angle Orthod*. 2004;74(1):31–36. doi:10.1043/0003-3219(2004)074<0031:EAOLLM>2.0.CO;2
6. Hang WM, Gelb M. Airway Centric® TMJ philosophy/Airway Centric® orthodontics ushers in the post-retraction world of orthodontics. *Cranio*. 2017;35(2):68–78. doi:10.1080/08869634.2016.1192315
7. Rogers AP. Muscle training and its relation to orthodontia. *Int J Orthod*. 1918;4(11):555–577. doi:10.1016/S1072-3471(18)80010-4
8. Singh G, ed. *Textbook of Orthodontics*. 3rd ed. New Delhi, India: Jaypee Brothers Medical Publisher; 2015. ISBN:978-93-5152-440-3.
9. Proffit WR. Equilibrium theory revisited: Factors influencing position of the teeth. *Angle Orthod*. 1978;48(3):175–186. doi:10.1043/0003-3219(1978)048<0175:ETRFIP>2.0.CO;2
10. Lipari MA, Pimentel G, Gamboa NA, Bayas I, Guerrero N, Miralles R. Electromyographic comparison of lips and jaw muscles between children with competent and incompetent lips: A cross-sectional study. *J Clin Pediatr Dent*. 2020;44(4):283–288. doi:10.17796/1053-4625-44.4.11
11. Stahl F, Grabowski R, Gaebel M, Kundt G. Relationship between occlusal findings and orofacial myofunctional status in primary and mixed dentition. Part II: Prevalence of orofacial dysfunctions. *J Orofac Orthop*. 2007;68(2):74–90. doi:10.1007/s00056-007-2606-9
12. Grabowski R, Kundt G, Stahl F. Interrelation between occlusal findings and orofacial myofunctional status in primary and mixed dentition. Part III: Interrelation between malocclusions and orofacial dysfunctions. *J Orofac Orthop*. 2007;68(6):462–476. doi:10.1007/s00056-007-0717-y
13. Miyamoto T, Yamada K, Hijiya K, et al. Ability to control directional lip-closing force during voluntary lip pursing in healthy young adults. *J Oral Rehabil*. 2019;46(6):526–532. doi:10.1111/joor.12776
14. Szyszka-Sommerfeld L, Sycińska-Dziarnowska M, Cernera M, Esposito L, Woźniak K, Spagnuolo G. Electromyographic assessment of muscle activity in children undergoing orthodontic treatment: A systematic review. *J Clin Med*. 2024;13(7):2051. doi:10.3390/jcm13072051
15. Woźniak K, Piątkowska D, Lipski M, Mehr K. Surface electromyography in orthodontics: A literature review. *Med Sci Monit*. 2013;19:416–423. doi:10.12659/MSM.883927
16. Ferrario VF, Tartaglia GM, Galletta A, Grassi GP, Sforza C. The influence of occlusion on jaw and neck muscle activity: A surface EMG study in healthy young adults. *J Oral Rehabil*. 2006;33(5):341–348. doi:10.1111/j.1365-2842.2005.01558.x
17. Castroflorio T, Bracco P, Farina D. Surface electromyography in the assessment of jaw elevator muscles. *J Oral Rehabil*. 2008;35(8):638–645. doi:10.1111/j.1365-2842.2008.01864.x

18. Hugger S, Schindler HJ, Kordass B, Hugger A. Clinical relevance of surface EMG of the masticatory muscles. (Part 1): Resting activity, maximal and submaximal voluntary contraction, symmetry of EMG activity. *Int J Comput Dent*. 2012;15(4):297–314. PMID:23457900.
19. Nishi SE, Basri R, Alam MK. Uses of electromyography in dentistry: An overview with meta-analysis. *Eur J Dent*. 2016;10(3):419–425. doi:10.4103/1305-7456.184156
20. Szyszka-Sommerfeld L, Woźniak K, Matthews-Brzozowska T, Kawala B, Mikulewicz M, Machoy M. The electrical activity of the masticatory muscles in children with cleft lip and palate. *Int J Paediatr Dent*. 2018;28(2):257–265. doi:10.1111/ipd.12349
21. De Felício CM, Sidequersky FV, Tartaglia GM, Sforza C. Electromyographic standardized indices in healthy Brazilian young adults and data reproducibility. *J Oral Rehabil*. 2009;36(8):577–583. doi:10.1111/j.1365-2842.2009.01970.x
22. Svensson P, Wang K, Sessle BJ, Arendt-Nielsen L. Associations between pain and neuromuscular activity in the human jaw and neck muscles. *Pain*. 2004;109(3):225–232. doi:10.1016/j.pain.2003.12.031
23. Hugger A, Hugger S, Schindler HJ. Surface electromyography of the masticatory muscles for application in dental practice: Current evidence and future developments. *Int J Comput Dent*. 2008;11(2):81–106. PMID:19119545.
24. Szyszka-Sommerfeld L, Machoy ME, Wilczyński S, Lipski M, Woźniak K. Superior orbicularis oris muscle activity in children surgically treated for bilateral complete cleft lip and palate. *J Clin Med*. 2021;10(8):1720. doi:10.3390/jcm10081720
25. Szyszka-Sommerfeld L, Sycińska-Dziarnowska M, Woźniak K, et al. The electrical activity of the orbicularis oris muscle in children with Down syndrome: A preliminary study. *J Clin Med*. 2021;10(23):5611. doi:10.3390/jcm10235611
26. Szyszka-Sommerfeld L, Woźniak K, Matthews-Brzozowska T, Kawala B, Mikulewicz M. Electromyographic analysis of superior orbicularis oris muscle function in children surgically treated for unilateral complete cleft lip and palate. *J Craniomaxillofac Surg*. 2017;45(9):1547–1551. doi:10.1016/j.jcms.2017.06.012
27. Tosello DO, Vitti M, Berzin F. EMG activity of the orbicularis oris and mentalis muscles in children with malocclusion, incompetent lips and atypical swallowing: Part II. *J Oral Rehabil*. 1999;26(8):644–649. doi:10.1046/j.1365-2842.1999.00409.x
28. Champely S. pwr: Basic Functions for Power Analysis_. R package version 1.3-0. Comprehensive R Archive Network (CRAN); 2020. https://CRAN.R-project.org/package=pwr. Accessed March 15, 2025.
29. 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://easystats.github.io/report/reference/report-package.html. Accessed March 15, 2025.
30. Segner D, Hasund A. *Individualisierte Kephalometrie*. 4th ed. Hamburg, Germany: Segner; 2003. ISBN:978-3-9802709-4-6.
31. Christensen LV, Hutchins MO. Methodological observations on positive and negative work (teeth grinding) by human jaw muscles. *J Oral Rehabil*. 1992;19(4):399–411. doi:10.1111/j.1365-2842.1992.tb01582.x
32. Soderberg GL, Knutson LM. A guide for use and interpretation of kinesiologic electromyographic data. *Phys Ther*. 2000;80(5):485–498. PMID:10792859.
33. Ambrosio AR, Trevilatto PC, Martins LP, Santos-Pinto AD, Shimizu RH. Electromyographic evaluation of the upper lip according to the breathing mode: A longitudinal study. *Braz Oral Res*. 2009;23(4):415–423. doi:10.1590/s1806-83242009000400011
34. Szyszka-Sommerfeld L, Sycińska-Dziarnowska M, Machoy M, et al. Electromyographic study of masticatory muscle function in children with Down syndrome. *J Clin Med*. 2022;11(3):506. doi:10.3390/jcm11030506
35. Smith DG, Clemens J, Crede W, Harvey M, Gracely EJ. Impact of multiple comparisons in randomized clinical trials. *Am J Med*. 1987;83(3):545–550. doi:10.1016/0002-9343(87)90768-6
36. Harradine NWT, Kirschen RHES. Lip and mentalis activity and its influence on incisor position: A quantitative electromyographic study. *Br J Orthod*. 1983;10(3):114–127. doi:10.1179/bjo.10.3.114
37. Lowe AA. Correlations between orofacial muscle activity and craniofacial morphology in a sample of control and anterior open-bite subjects. *Am J Orthod*. 1980;78(1):89–98. doi:10.1016/0002-9416(80)90042-1
38. Leonardo SE, Sato Y, Kaneko T, Yamamoto T, Handa K, Iida J. Differences in dento-facial morphology in lip competence and lip incompetence. *Orthod Waves*. 2009;68(1):12–19. doi:10.1016/j.odw.2008.11.002
39. Störmer K, Pancherz H. Electromyography of the perioral and masticatory muscles in orthodontic patients with atypical swallowing. *J Orofac Orthop*. 1999;60(1):13–23. doi:10.1007/bf01358712
40. López-Soto LM, López-Soto OP, Osorio-Forero A, Restrepo F, Tamayo-Orrego L. Muscle activity and muscle strength in atypical swallowing. *Rev Salud Uninorte*. 2017;33:273–284. https://www.redalyc.org/pdf/817/81753881002.pdf. Accessed March 15, 2025.
41. Ahlgren J. EMG studies of lip and cheek activity in sucking habits. *Swed Dent J*. 1995;19(3):95–101. PMID:7676390.
42. Regalo SCH, Vitti M, Moraes MTB, et al. Electromyographic analysis of the orbicularis oris muscle in oralized deaf individuals. *Braz Dent J*. 2005;16(3):237–242. doi:10.1590/s0103-64402005000300012
43. Martinelli RLC, Marchesan IQ, Gusmão RJ, Berretin-Felix G. Effect of lingual frenotomy on tongue and lip rest position: A nonrandomized clinical trial. *Int Arch Otorhinolaryngol*. 2022;26(1):e069–e074. doi:10.1055/s-0041-1726050
44. Santos HKMPDS, Cunha DAD, Andrade RAD, et al. Effects of lingual frenotomy on breastfeeding and electrical activity of the masseter and suprahyoid muscles. *Codas*. 2023;35(2):e20210262. doi:10.1590/2317-1782/20232021262
45. Saccucci M, Tecco S, Ierardoa G, Luzzi V, Festa F, Polimeni A. Effects of interceptive orthodontics on orbicular muscle activity: A surface electromyographic study in children. *J Electromyogr Kinesio*. 2011;21(4):665–671. doi:10.1016/j.jelekin.2011.03.005
46. Germer CM, Farina D, Elias LA, Nuccio S, Hug F, Del Vecchio A. Surface EMG cross talk quantified at the motor unit population level for muscles of the hand, thigh, and calf. *J Appl Physiol*. 2021;131(2):808–820. doi:10.1152/japplphysiol.01041.2020
47. De Luca CJ, Kuznetsov M, Gilmore LD, Roy SH. Inter-electrode spacing of surface EMG sensors: Reduction of crosstalk contamination during voluntary contractions. *J Biomech*. 2012;45(3):555–561. doi:10.1016/j.jbiomech.2011.11.010
48. Bianchini AP, Guedes ZCF, Vieira MM. A study on the relationship between mouth breathing and facial morphological pattern. *Braz J Otorhinolaryngol*. 2007;73(4):500–505. doi:10.1016/s1808-8694(15)30101-4
49. Yang X, Lai G, Wang J. Effect of orofacial myofunctional therapy along with preformed appliances on patients with mixed dentition and lip incompetence. *BMC Oral Health*. 2022;22(1):586. doi:10.1186/s12903-022-02645-w
50. Hassan AH, Turkistani AA, Hassan MH. Skeletal and dental characteristics of subjects with incompetent lips. *Saudi Med J*. 2014;35(8):849–854. PMID:25129185.
52. Rasheed M, Sajjad B, Devi A, Shaikh WG, Naureen H, Azfar M. Comparison of mean upper lip length in individuals with competent lips, lips apart and incompetent lips. *Pak J Med Health Sci*. 2023;17(2):286–288. doi:10.53350/pjmhs2023172286
52. Takada JI, Miyamoto JJ, Sato C, Dei A, Moriyama K. Comparison of EMG activity and blood flow during graded exertion in the orbicularis oris muscle of adult subjects with and without lip incompetence: A cross-sectional survey. *Eur J Orthod*. 2018;40(3):304–311. doi:10.1093/ejo/cjx061
53. Kayal V, Nikhil DKR. Prevalence of lip incompetence in 6 to 12 year-old children visiting university hospital: A retrospective study: Original Research. *Int J Pedodont Rehabil*. 2024;9(2):26–34. doi:10.56501/intjpedorehab.v9i2.1140
54. Yoshizawa S, Ohtsuka M, Kaneko T, Iida J. Study of training for improving lip incompetence. *Orthod Waves*. 2016;75(3):47–53. doi:10.1016/j.odw.2016.07.001
55. Alizade AS, Asadi E, Jafari-Naeimi A, Kalbassi S. Efficacy of the combination of myofunctional therapy (lip exercises) and activator high-pull headgear in the closure of interlabial gap in long-face skeletal class II patients with lip incompetence: A 6–8-month longitudinal randomized clinical trial. *Dent Res J (Isfahan)*. 2024;21:3. PMID:38425317. PMCID:PMC10899158.
56. Fang ML, Choi SH, Choi YJ, Lee KJ. Pattern of lip retraction according to the presence of lip incompetence in patients with class II malocclusion. *Korean J Orthod*. 2023;53(4):276–285. doi:10.4041/kjod22.260

TRPC3 induces intervertebral disc degeneration by mediating the Ca^{2+}/NF-κB pathway to inhibit autophagy

Yingchao Gao[1,A,D–F], Ning Zhang[2,B,C,F], Jun-Fei Zhang[3,B,C,F], Zhengqi Fei[1,A,E,F]

[1] Department of Orthopedics, 942nd Hospital of the Joint Logistics Support Force of the People's Liberation Army of China, Yinchuan, China
[2] Department of Orthopedics, General Hospital of Ningxia Medical University, Yinchuan, China
[3] School of Clinical Medicine, Ningxia Medical University, Yinchuan, 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(5):819–833

Address for correspondence
Yingchao Gao
E-mail: jirui.999@163.com

Funding sources
This study was supported by the Ningxia Natural Science Foundation (grant No. 2021AAC03413) and the Joint Logistic Support Force 942 Hospital Research Program (grant No. D942YYKY0003).

Conflict of interest
None declared

Received on January 20, 2025
Reviewed on May 18, 2025
Accepted on June 25, 2025

Published online on May 5, 2026

Abstract

Background. Intervertebral disc degeneration (IDD) is the primary cause of lower back pain. Transient receptor potential canonical 3 (*TRPC3*) is a nonselective cation channel permeable to Ca^{2+}.

Objectives. This study explores the mechanisms by which the *TRPC3*-mediated Ca^{2+}/nuclear factor kappa B (NF-κB) pathway regulates autophagy in IDD.

Materials and methods. An IDD rat model was established using the annulus fibrosus puncture method and was treated with local intraspinal injection of adeno-associated virus (AAV)-shRNA targeting *TRPC3*. Primary human nucleus pulposus cells (NPCs) were transfected with *TRPC3* siRNA and subsequently treated with pyrrolidine dithiocarbamate (PDTC; an NF-κB inhibitor), rapamycin (RAPA), or 3-methyladenine (3-MA), respectively. Micro-computed tomography (micro-CT), hematoxylin and eosin (H&E) staining, immunohistochemistry, western blotting, transmission electron microscopy (TEM), and flow cytometry were performed.

Results. *TRPC3* expression was significantly increased in IDD rats ($p < 0.05$). *TRPC3* shRNA ameliorated histopathological damage in IDD rats and promoted the expression of autophagy-related protein 5 (*ATG5*), *Beclin-1*, and *LC3-II* (all $p < 0.05$). In vitro, interleukin-1 beta (*IL-1β*) increased Ca^{2+} levels, siRNA *TRPC3* reduced them, and PDTC further decreased them ($p < 0.05$). In addition, siRNA *TRPC3* increased the expression of *ATG5*, *Beclin-1*, and the *LC3-II/LC3-I* ratio and inhibited phosphorylation of *p*-NF-κB *p65* in NPCs ($p < 0.05$). Transmission electron microscopy and flow cytometry showed that siRNA *TRPC3*-induced autophagy promoted apoptosis in NPCs ($p < 0.05$). Furthermore, siRNA *TRPC3* increased the levels of aggrecan and collagen II and decreased matrix metalloproteinase-13 (*MMP-13*) expression ($p < 0.05$).

Conclusions. *TRPC3* exacerbates IDD by inhibiting protective autophagy via activation of the Ca^{2+}/NF-κB signaling pathway. Knockdown of *TRPC3* promotes autophagy, which in turn influences NPC apoptosis and extracellular matrix (ECM) metabolism. This study offers potential novel strategies for IDD prevention and treatment.

Key words: autophagy, NF-κB, intervertebral disc degeneration, *TRPC3*, calcium ions influx

Cite as
Gao Y, Zhang N, Zhang JF, Fei Z. *TRPC3* induces intervertebral disc degeneration by mediating the Ca^{2+}/NF-κB pathway to inhibit autophagy. *Adv Clin Exp Med*. 2026;35(5):819–833. doi:10.17219/acem/207572

DOI
10.17219/acem/207572

Highlights

- *TRPC3* is upregulated in intervertebral disc degeneration (IDD), driving pathological Ca^{2+} influx and activation of the NF-κB signaling pathway in nucleus pulposus cells.
- Silencing *TRPC3* restores autophagy in IDD, significantly increasing *ATG5*, *Beclin-1*, and *LC3-II* expression while reducing NF-κB p65 phosphorylation.
- *TRPC3* inhibition improves disc matrix homeostasis, enhancing aggrecan and collagen II levels and suppressing *MMP-13*-mediated extracellular matrix (ECM) degradation.
- Targeting the *TRPC3*/Ca^{2+}/NF-κB axis represents a novel therapeutic strategy for preventing and treating IDD.

Background

Low back pain (LBP) is the most common chronic musculoskeletal condition affecting adults and has emerged as a leading cause of global disability, profoundly impacting patients' work capacity and daily life.[1] Research indicates that 60–80% of individuals experience symptoms of LBP, with approx. 10% of patients becoming disabled.[2–4] Currently, approx. 630 million people worldwide suffer from neck pain and LBP. Intervertebral disc degeneration (IDD) is the primary cause of LBP and disc-related disorders, such as herniation and spinal stenosis.[5] Studies have revealed that approx. 40% of low back and leg pain cases are attributable to IDD.[6] The prevalence of IDD is steadily increasing across all age groups, posing significant challenges for medical care and society.[7] Although IDD is particularly prevalent among the elderly population, a shift toward younger patients has been observed in recent years, likely due to changes in lifestyle and habits.[8] Intervertebral disc degeneration is a progressive, multifactorial condition characterized by biomechanical, structural, and biological changes in disc tissue induced by various factors, ultimately leading to the loss of disc integrity and function. These changes include annulus fibrosus rupture, nucleus pulposus (NP) herniation, extracellular matrix (ECM) degradation, reduced disc height, and compression of the spinal cord and nerve roots, ultimately resulting in lower back and leg pain.[9] Recent research suggests that IDD is a complex disease resulting from the interaction of multiple factors, including aging, genetic predisposition, mechanical stress, inflammation, oxidative stress, metabolic dysfunction, environmental factors, and autophagy.[5,10–12] Furthermore, IDD is characterized by degeneration of the intervertebral disc ECM, leading to reduced biomechanical integrity and pain.[13] Among these factors, matrix metalloproteinases (MMPs) are key contributors to ECM degradation in IDD. Previous studies have shown that bovine bone grafts can contribute to intervertebral disc repair by supporting cell attachment, proliferation, and matrix synthesis.[14] Recent studies have also demonstrated that inhibition of matrix degradation can reduce mitochondrial damage, inhibit cell apoptosis and senescence, and significantly delay the progression of IDD.[15] However, the pathogenesis of IDD remains a major challenge in clinical practice. Therefore, it is crucial to further elucidate the mechanisms underlying IDD and identify reliable therapeutic targets to improve ECM metabolic imbalance and ultimately reverse the progression of IDD.

As an intracellular recycling process, autophagy degrades cytoplasmic components via lysosomes and plays a vital role in cellular self-degradation and recycling.[16] This process is essential for maintaining metabolic homeostasis. It has been reported that autophagy alleviates osteoarthritis by regulating ECM metabolism.[17] Recent studies have increasingly demonstrated that autophagy plays a crucial role in IDD.[18,19] Cheng et al. has found that regulation of chaperone-mediated autophagy can effectively delay the progression of IDD in an inflammatory environment.[20] However, the role of autophagy in IDD remains controversial because of its dual effects. Some studies have reported that the expression levels of autophagy-associated genes are significantly higher in IDD than in healthy discs.[21] In addition, research has indicated that appropriate activation of autophagy can protect nucleus pulposus cells (NPCs), which are the core cellular components of the intervertebral disc, from pressure-induced damage. In contrast, impairment of autophagic function can lead to apoptosis of NPCs, thereby further accelerating the progression of IDD.[22,23]

Transient receptor potential (TRP) channels are known for their involvement in sensory processes, such as temperature and pain perception, as well as in the regulation of cellular calcium homeostasis. Transient receptor potential canonical 3 (*TRPC3*) is a Ca^{2+}-permeable, nonselective cation channel that plays vital roles in a wide range of cellular physiological processes. *TRPC3* facilitates calcium entry, thereby enabling cells to regulate gene expression, as well as cell growth and differentiation. Notably, *TRPC3* has been implicated in various physiological and pathological conditions, including cardiovascular diseases, neurological disorders, and cancer. In the context of IDD, the specific mechanisms of *TRPC3* remain controversial. Some studies suggest that *TRPC3* is involved in cytosolic Ca^{2+} elevation, activation of nuclear factor kappa B (NF-κB), and cytokine upregulation.[24] Notably, research has shown that increased *TRPC3* expression can lead to elevated intracellular Ca^{2+} concentrations and is associated with decreased bone mass.[25] Abnormal *TRPC3* channel activity may promote bone resorption, reduce bone

density, and mediate the Ca^{2+}/NF-κB signaling pathway.[26] Calcium (Ca^{2+}) is one of the most abundant and important signaling molecules in the human body. In tissues such as intervertebral discs, bone, and cartilage, elevated intracellular Ca^{2+} concentrations activate Ca^{2+} signaling pathways, which in turn regulate gene expression and protein synthesis, thereby influencing changes in the microenvironment of intervertebral disc tissue. Previous investigations have demonstrated the critical role of Ca^{2+} signaling in intervertebral disc tissues. Intervertebral disc degeneration leads to alterations in the osmotic pressure of intervertebral disc tissue, resulting in activation of the Ca^{2+} signaling pathway. Intervertebral disc cells regulate gene expression and protein synthesis by increasing intracellular Ca^{2+} concentrations. Furthermore, Ca^{2+} acts as a second messenger in the human body, activating downstream signaling pathways. The role of the NF-κB signaling pathway in IDD has been widely recognized; however, its upstream regulatory mechanisms are diverse. Evidence suggests that calcium/calmodulin-dependent protein kinase II, in conjunction with interleukin-1 receptor-associated kinase 1 (*IRAK1*), plays a critical role in the phosphorylation processes that activate NF-κB.[27] When activated, the NF-κB pathway triggers cell apoptosis and ECM degradation, thereby promoting IDD.[28] In addition, the NF-κB pathway promotes autophagy in various diseases. Melatonin has been shown to induce autophagy via the NF-κB signaling pathway, thereby preventing ECM degeneration in intervertebral disc cells.[29] Furthermore, through the *miR-139-3p*/*CXCR4*/*NF-κB* axis, *lncRNA H19* enhances autophagy and apoptosis in NPCs, thereby aggravating IDD.[30] However, there are currently no reports on whether *TRPC3* can activate autophagy to regulate IDD via the Ca^{2+}/NF-κB pathway. Moreover, whether *TRPC3*-induced NF-κB activation through Ca^{2+} influx exerts synergistic effects on apoptosis and ECM degradation remains to be further elucidated.

Objectives

This study aims to investigate the role and underlying mechanisms by which the *TRPC3*-mediated Ca^{2+}/NF-κB pathway inhibits autophagy in IDD using both in vivo and in vitro approaches.

Materials and methods

Animals

Twenty-four specific pathogen-free (SPF) healthy male Sprague Dawley (SD) rats (8 weeks old; body weight, 200 ±20 g) were purchased from Chengdu Dashuo Biotechnology Co., Ltd. (Chengdu, China). The rats were housed in a pathogen-free environment for 1 week to acclimate to the laboratory conditions, with free access to a standard diet (caloric composition: 70% carbohydrate, 10% fat, and 24% protein; provided by Chengdu Dashuo Biotechnology Co., Ltd.) and water, under a normal 12-h light/dark cycle.

Experimental design

The rats were randomly assigned to 4 groups: the sham group (n = 6), model group (n = 6), shRNA negative control (NC) group (n = 6), and shRNA *TRPC3* group (n = 6). An IDD rat model was established using the annulus fibrosus puncture method.[31] Twelve hours before model induction, the rats were fasted with ad libitum access to water. The rats were weighed and then intraperitoneally injected with 40 mg/kg of 1% pentobarbital sodium for anesthesia, based on body weight. Once anesthetized, the rats were placed in the supine position with their limbs immobilized. The surgical area was shaved and disinfected with iodine. A longitudinal incision of 3–4 cm was made approx. 0.5 cm to the right of the midline. Each tissue layer was sequentially incised to expose the posterior abdominal wall. The intestinal tract and greater omentum were gently retracted to prevent injury, and the paravertebral lumbar muscles were bluntly dissected. The L4/5 and L5/6 intervertebral discs were exposed, and a 21-gauge needle was used to puncture the annulus fibrosus at 3 points on the right anterior aspect of the vertebral body, directed toward the center of the intervertebral disc. The puncture depth was approx. 2–3 mm, limiting the injury to the full thickness of the annulus fibrosus. Following the injury, the needle was maintained in position for 10 s. After successful puncture, the incision was closed in layers. Rats in the sham surgery group underwent the same skin and tissue incision without intervertebral disc puncture. Postoperatively, routine care was provided, including intramuscular injection of penicillin (8×10^5 U/day, once daily for 3 consecutive days) to prevent infection. The rats were closely monitored postoperatively for food intake, gait, wound infection, and urinary retention.

Three rats were randomly selected from the sham and model groups at 8 weeks after surgery. The rats were anesthetized with isoflurane, followed by cervical dislocation at the atlantoaxial joint. Tissue samples were collected, and the skin, paravertebral muscles, and ligaments were dissected layer by layer. The L5/6 intervertebral disc tissue was rapidly harvested and fixed in 4% paraformaldehyde. To evaluate the success of model establishment and observe morphological changes, hematoxylin and eosin (H&E) staining was performed.

Model establishment was confirmed by pathological examination at 8 weeks after surgery. On the 1st day after successful model establishment, rats in the shRNA NC group received a local intraspinal injection of adeno-associated virus (AAV)-shRNA NC (1×10^{12} μg/kg). Rats in the shRNA *TRPC3* group received a local intraspinal injection of AAV-shRNA *TRPC3* (1×10^{12} μg/kg). Both AAV-shRNA NC and AAV-shRNA *TRPC3* were designed and synthesized by Shanghai Jikai Gene Medical Technology Co., Ltd.

(Shanghai, China). After 3 weeks, the rats were anesthetized with an overdose of isoflurane, and NP tissue samples were collected from the intervertebral discs of each group.

H&E staining

Intervertebral disc tissues were first decalcified using a 15% ethylenediaminetetraacetic acid (EDTA) solution. The tissues were then dehydrated and embedded, sectioned at a thickness of 5 μm, stained with hematoxylin for 10–20 min followed by eosin for 3–5 min, and mounted with neutral gum.[32] Images were acquired using a Pannoramic 250 FLASH III digital slide scanner (3DHISTECH, Budapest, Hungary).

Immunohistochemistry staining

Intervertebral disc tissues were sectioned, deparaffinized, and rehydrated. The sections were then heated in 10 mM sodium citrate buffer (pH 6.0) at 100°C for 30 min for antigen retrieval and blocked with 3% bovine serum albumin (BSA; cat. No. GC305010; Servicebio, Wuhan, China) for 1 h at room temperature. The sections were subsequently incubated overnight at 4°C with primary antibodies against autophagy-related protein 5 (*ATG5*; bs-4005R, 1:100; Proteintech, Wuhan, China), *Beclin-1* (11306-1-AP, 1:100, Proteintech), and *LC3-II* (14600-1-AP, 1:200; Proteintech). Afterward, the sections were incubated with a secondary antibody (horseradish peroxidase (HRP)-labeled goat anti-rabbit immunoglobulin G (IgG; H+L) GB23303, 1:100; Servicebio) for 1 h at room temperature. Diaminobenzidine (DAB) color development and hematoxylin counterstaining were then performed.[33] Finally, the sections were examined under a light microscope (BA400Digital; Motic China Group Co., Ltd., Xiamen, China), and images were captured using a digital trinocular microscope (BA400 Digital; McAudi, Xiamen, China).

Micro-computed tomography scan

Intervertebral disc tissue samples were scanned using the Micro-CT Scanner software. For three-dimensional analysis, the samples were analyzed using CTAn software (Bruker Micro-CT, Billerica, USA) to obtain the following parameters: structure model index (SMI), trabecular thickness (Tb.Th), bone volume to total volume ratio (BV/TV), trabecular number (Tb.N), and trabecular separation (Tb.Sp).

Primary human NPCs isolation and culture

Nucleus pulposus tissue was obtained during lumbar discectomy, and its morphology was observed under a light microscope (Leica DMI1; Leica Microsystems, Wetzlar, Germany). Following washing with phosphate-buffered saline (PBS), NP tissue samples were digested with 0.25% trypsin and 0.2% collagenase type II at 37°C for 4–6 h. After removal of tissue debris by filtration through a 200-μm filter, purified NPCs were cultured in Dulbecco's modified Eagle's medium (DMEM)/F-12 supplemented with 10% fetal bovine serum (FBS), 100 μg/mL streptomycin, and 100 μg/mL penicillin at 37°C in a humidified atmosphere containing 5% CO_2. Nucleus pulposus cells at passage 2 were used for subsequent in vitro experiments.

Cell transfection and grouping

The cell experiments were divided into 2 parts.

Part I: NPCs were divided into the following groups: control, IL-1β, IL-1β + siRNA NC, IL-1β + siRNA *TRPC3*, and IL-1β + siRNA *TRPC3* + pyrrolidine dithiocarbamate (PDTC; NF-κB inhibitor, 100 μmol/L).

Part II: NPCs were divided into the following groups: control, IL-1β, IL-1β + siRNA NC, IL-1β + siRNA *TRPC3*, IL-1β + siRNA *TRPC3* + 3-methyladenine (3-MA; autophagy inhibitor), and IL-1β + siRNA *TRPC3* + rapamycin (RAPA; autophagy activator).

Cell transfection was performed using the RiboFect™ CP Transfection Kit (C10511-05; RiboBio, Guangzhou, China). The lyophilized siRNA was reconstituted in RNase-free water to obtain a 20 μM stock solution. The transfection mixture was prepared by thoroughly mixing 120 μL of RiboFect™ CP Buffer, 12 μL of RiboFect™ CP Reagent, and 10 μL of siRNA, followed by incubation at 37°C. Except for the control group, cells in all other groups were treated with 10 ng/mL IL-1β after cell adherence to induce cellular injury. Simultaneously, cells in the IL-1β + siRNA *TRPC3* + 3-MA group were treated with 3-MA (5 mmol/L), and cells in the IL-1β + siRNA *TRPC3* + RAPA group were treated with RAPA (250 nmol/L). The corresponding assays were performed 24 h later.

Western blot

Rat intervertebral disc NP tissue or NPCs were lysed for 10 min in radioimmunoprecipitation assay (RIPA) lysis buffer (Beyotime, Shanghai, China). Protein concentrations were quantified using a bicinchoninic acid (BCA) Protein Assay Kit (Beyotime). Protein samples from the supernatant were mixed with an equal volume of sodium dodecyl sulfate (SDS) loading buffer and boiled for 5 min to denature the proteins. Equal amounts of protein were loaded onto a 12% polyacrylamide gel and subjected to electrophoresis for 60–90 min. Subsequently, proteins were transferred onto polyvinylidene fluoride (PVDF) membranes (Merck Millipore, Billerica, USA). To block nonspecific binding, the membranes were incubated with 5% skim milk for 1 h at room temperature. The PVDF membranes were then incubated with primary antibodies overnight at 4°C, followed by incubation with secondary antibodies for 2 h at room temperature.

The antibodies used were as follows: anti-β-actin (cat. No. AC026, 1:50,000), anti-*ATG5* (cat. No. A0203, 1:1,000), anti-*Beclin-1* (cat. No. A7353, 1:2,000), anti-*LC3B* (cat. No. A19665, 1:2,000), anti-*NF-κB* p65 (cat. No. A2547, 1:2,000), anti-phospho-*NF-κB* p65 (cat. No. AP0123, 1:2,000), and anti-*MMP13* (cat. No. A11148, 1:2,000), all purchased from ABclonal Biotechnology Co., Ltd. (Wuhan, China); anti-*TRPC3* (cat. No. 77934, 1:1,000) was purchased from Cell Signaling Technology (CST; Danvers, USA); anti-aggrecan (cat. No. DF7561, 1:1,000) and goat anti-rabbit IgG (H+L) HRP (cat. No. S0001, 1:5,000) were purchased from Affinity Biosciences (Beijing, China); and anti-collagen II (cat. No. BS-10589R, 1:2,000) was purchased from Bioss Biotechnology Co., Ltd. (Beijing, China).

Protein bands were visualized using an enhanced chemiluminescence (ECL) kit (Biosharp, Hefei, China), and signals were captured using the Tanon 5200 Multi-System (Tanon, Shanghai, China).

Transmission electron microscope for autophagy observation

A transmission electron microscope (TEM) was used to investigate autophagy in NPCs. First, the NPC samples were subjected to primary fixation with 3% glutaraldehyde to preserve cellular structures. Subsequently, secondary fixation was performed using 1% osmium tetroxide to enhance contrast of the cellular components. Ultrathin sections with a thickness of approximately 60 nm were prepared using an ultramicrotome (Leica Camera AG). To improve electron density and visualization of cellular organelles, the sections were stained with uranyl acetate for 15 min. Subsequently, the sections were briefly stained with lead citrate for 2 min to further enhance contrast. Finally, the stained samples were examined using a JEM-1400 FLASH transmission electron microscope (JEOL, Tokyo, Japan).

Detection of cell apoptosis by flow cytometry

To assess the apoptotic status of NPCs, flow cytometric analysis was performed using the Annexin V–APC/PI Apoptosis Detection Kit (KGA1030; KeyGen Biotech Corp., Ltd., Nanjing, China) according to the manufacturer's instructions.[34] Nucleus pulposus cells were seeded in 6-well plates at a density of 2×10^5 cells per well and incubated at 37°C in a humidified atmosphere containing 5% CO_2 to allow cell attachment and growth. Once the cells reached appropriate confluency, apoptosis detection was initiated. A total of 5 µL of Annexin V–FITC and 5 µL of propidium iodide (PI) were added to the cell suspension. The cells were then incubated for 15 min at room temperature in the dark to allow effective binding of the dyes to the cell membranes. After incubation, the stained NPCs were analyzed using a flow cytometer (CytoFLEX; Beckman Coulter, Brea, USA).

Statistical analyses

For statistical analysis, IBM SPSS Statistics v. 25 (IBM Corp., Armonk, USA) was used, and data are presented as the median (minimum–maximum). The Kruskal–Wallis (K–W) test, followed by Dunn's post hoc test with Bonferroni correction, was applied for multiple-group comparisons. A $p < 0.05$ was considered statistically significant. Statistical analysis results are presented in Supplementary Tables 1–6.

Results

Inhibition of *TRPC3* improved histopathological damage in IDD rats

To investigate the role and mechanisms of *TRPC3* in IDD, AAV-shRNA *TRPC3* was injected into the intervertebral disc tissue of rats in the animal study. First, intervertebral disc tissues were evaluated using micro-CT, as shown in Table 1. Compared with the sham group, the model group exhibited decreased BV/TV, Tb.N, and Tb.Th values, while Tb.Sp and SMI were increased; however, only BV/TV, Tb.N, and SMI showed statistically significant differences (all $p = 0.050$). However, no significant difference was observed between the shRNA *TRPC3* group and the shRNA NC group ($p = 0.127$, $p = 0.513$, $p = 0.275$). To further evaluate histopathological changes, H&E staining was performed. In the model group, a significant reduction in NPCs was observed compared with the sham group, whereas chondrocytes and the cartilaginous matrix were increased and arranged in clusters. The shRNA NC group did not exhibit notable improvement in pathological conditions compared with the model group. In contrast, the shRNA *TRPC3* group exhibited a relatively intact intervertebral disc tissue structure: the annulus fibrosus was arranged in concentric circles with multiple layers of fibrocartilage and showed a denser organization, and the NP was rich in elastic gel-like material, with abundant NPCs and ECM, indicating attenuation of pathological

Table 1. Microstructural parameters of intervertebral disc tissue analyzed with micro-CT

Group	Sham	Model	shRNA NC	shRNA *TRPC3*
BV/TV [%]	37.44 ±1.67	30.74 ±1.95#	30.17 ±2.59	32.49 ±1.25
Tb.N [mm]	3.51 ±0.07	2.93 ±0.24#	2.92 ±0.22	3.07 ±0.04
Tb.Th [mm]	0.11 ±0.002	0.10 ±0.003	0.10 ±0.005	0.11 ±0.013
Tb.Sp [mm]	0.25 ±0.01	0.28 ±0.05	0.28 ±0.02	0.26 ±0.01
SMI	0.21 ±0.09	0.82 ±0.14#	0.79 ±0.17	0.64 ±0.12

CT – computed tomography; BV/TV – percent bone volume; Tb.Th – trabecular thickness; Tb.Sp – trabecular separation; Tb.N – trabecular number; SMI – structure model index. Data were analyzed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction; #$p < 0.05$, compared with the sham group.

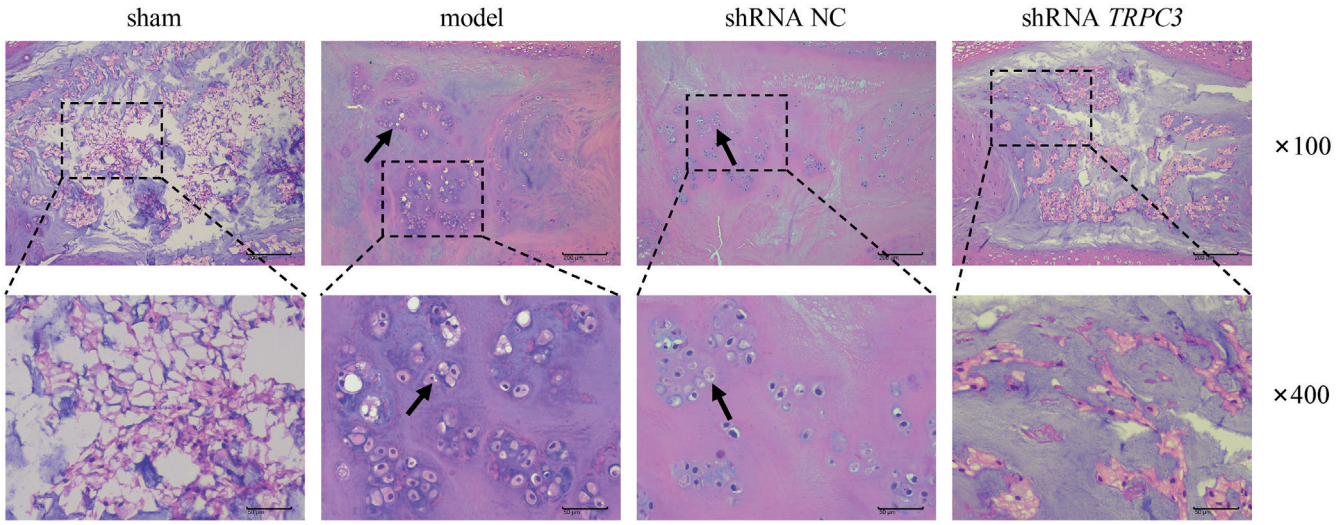

Fig. 1. Inhibition of *TRPC3* ameliorates histopathological damage in intervertebral disc degeneration (IDD) rats. Hematoxylin and eosin (H&E) staining was used to observe pathological changes in intervertebral disc tissues. Scale bars = 50 μm (×400) and 200 μm (×100). Black arrows indicate reduced numbers of NPCs and increased chondrocytes and cartilaginous matrix

TRPC3 – transient receptor potential canonical 3; NC – negative control; NPCs – nucleus pulposus cells.

damage (Fig. 1). Therefore, these results indicate that *TRPC3* knockdown can ameliorate histopathological damage in IDD rats in vivo.

Inhibition of *TRPC3* promoted autophagy in the nucleus pulposus of intervertebral discs in vivo

Subsequently, we analyzed the effect of shRNA *TRPC3* on autophagy levels in the NP of intervertebral discs in IDD rats using IHC staining. The expression levels of *ATG5*, *Beclin-1*, and *LC3-II* were markedly higher in the model group than in the sham group (all p = 0.050). Moreover, treatment with AAV-shRNA *TRPC3* further increased the expression of *ATG5*, *Beclin-1*, and *LC3-II* compared with the shRNA NC group (all p = 0.050) (Fig. 2, Supplementary Table 1). In addition, the model group exhibited a marked increase in *TRPC3* expression in NP tissue compared with the sham group, as confirmed by western blot analysis (p = 0.05). In contrast, *TRPC3* expression levels were significantly decreased in the AAV-shRNA *TRPC3*-treated group relative to the shRNA NC group (p = 0.024) (Fig. 3, Supplementary Table 2). Overall, these findings suggest that high *TRPC3* expression may lead to abnormal accumulation of autophagosomes by inhibiting autophagy, thereby aggravating IDD, whereas *TRPC3* knockdown may reverse the IDD phenotype by activating protective autophagy signaling.

The knockdown of *TRPC3* inhibited the Ca^{2+}/NF-κB pathway to promote autophagy in NPCs in vitro

To further investigate the mechanisms by which *TRPC3* regulates autophagy in IDD, in vitro experiments were performed using primary human NPCs. As shown in Fig. 4A, intracellular Ca^{2+} levels were significantly higher in the IL-1β-treated group than in the control group (p = 0.003). Following siRNA-mediated knockdown of *TRPC3*, Ca^{2+} levels were significantly reduced compared with those in the shRNA NC group (p = 0.050). In addition, treatment of NPCs with PDTC, a selective NF-κB inhibitor, led to a further reduction in intracellular Ca^{2+} levels (p = 0.050). Western blot analysis showed that knockdown of *TRPC3* attenuated the IL-1β-induced increase in phosphorylated NF-κB p65 (p = 0.011), while total NF-κB p65 expression remained unchanged (p = 0.273) (Fig. 4B–D). Notably, PDTC treatment significantly suppressed p-NF-κB p65 phosphorylation in the presence of siRNA *TRPC3* (p = 0.050) (Fig. 4D, Supplementary Table 3). Furthermore, siRNA-mediated knockdown of *TRPC3* increased the expression of *ATG5*, *Beclin-1*, and the *LC3-II/LC3-I* ratio compared with the IL-1β + siRNA NC group (p < 0.05). Moreover, NPCs treated with PDTC in combination with siRNA *TRPC3* and IL-1β exhibited further increases in these autophagy-related proteins (all p = 0.050) (Fig. 5A–D). In addition, PDTC significantly suppressed the IL-1β-induced increase in *TRPC3* expression (p = 0.050) (Fig. 5E, Supplementary Table 4). Collectively, these results indicate that *TRPC3* activates the NF-κB pathway through Ca^{2+} influx and promotes autophagy-related protein expression, ultimately exacerbating IDD, whereas combined inhibition of *TRPC3* and NF-κB exerts a synergistic effect in enhancing protective autophagy, thereby delaying IDD progression.

The knockdown of *TRPC3*-activated autophagy promoted apoptosis in NPCs

To further examine the impact of *TRPC3*-induced autophagy on cell apoptosis, NPCs were treated with rapamycin

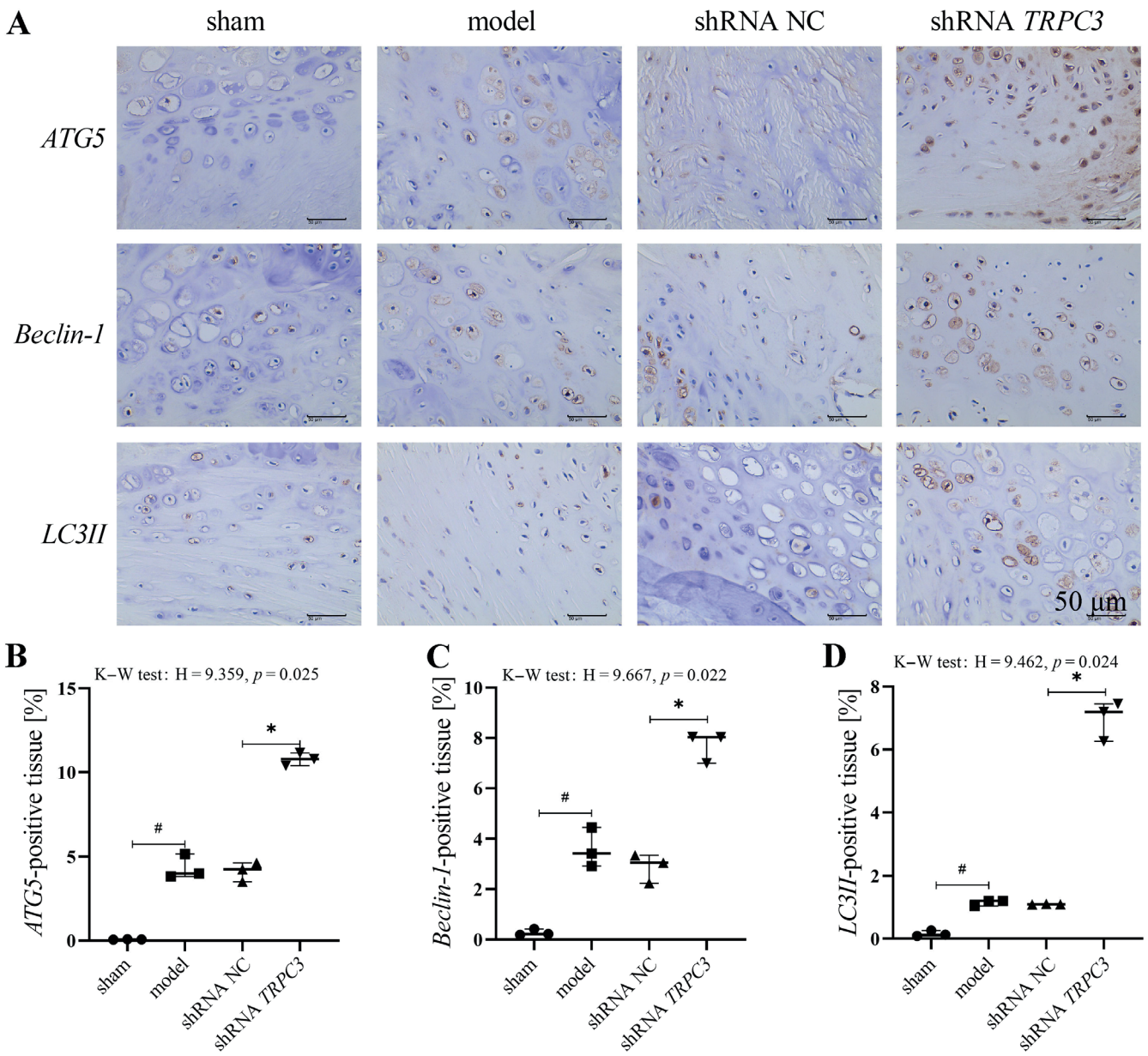

Fig. 2. Inhibition of *TRPC3* promotes autophagy-related protein expression in the nucleus pulposus of intervertebral discs. A. Immunohistochemical staining of *ATG5*, *Beclin-1*, and *LC3-II* in intervertebral disc tissues. Scale bar = 50 μm. Hematoxylin staining shows cell nuclei in blue, and diaminobenzidine (DAB) staining indicates positive expression in brown; B–D. Quantitative analysis of the positive area ratios of *ATG5*, *Beclin-1*, and *LC3-II* immunohistochemical staining. Data are presented as the median (minimum–maximum) (n = 3). Data were analyzed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction

#p < 0.05 vs the sham group; *p < 0.05 vs the shRNA negative control (NC) group; *TRPC3* – transient receptor potential canonical 3; DAB – diaminobenzidine; *ATG5* – autophagy-related protein 5.

(RAPA) and 3-methyladenine (3-MA). In the IL-1β group, TEM revealed the presence of autophagosomes and primary lysosomes, whereas the IL-1β + siRNA *TRPC3* group exhibited a markedly increased number of autophagosomes and primary lysosomes. Treatment with RAPA further enhanced autophagy, whereas 3-MA reduced autophagic activity (Fig. 6). Furthermore, compared with the control group, IL-1β treatment markedly increased NPC apoptosis (p = 0.050). Following siRNA-mediated knockdown of *TRPC3*, NPC apoptosis was significantly higher than that in the shRNA NC group (p = 0.050). Compared with the IL-1β + siRNA *TRPC3* group, treatment with RAPA resulted in a further increase in NPC apoptosis (p = 0.039), whereas 3-MA led to a reduction in apoptosis (Fig. 7, Supplementary Table 5).

These results suggest that *TRPC3* alleviates IL-1β-induced apoptosis in NPCs by inhibiting autophagy, whereas excessive activation of autophagy exacerbates cell death, indicating a bidirectional role of the *TRPC3*-autophagy axis in the regulation of apoptosis.

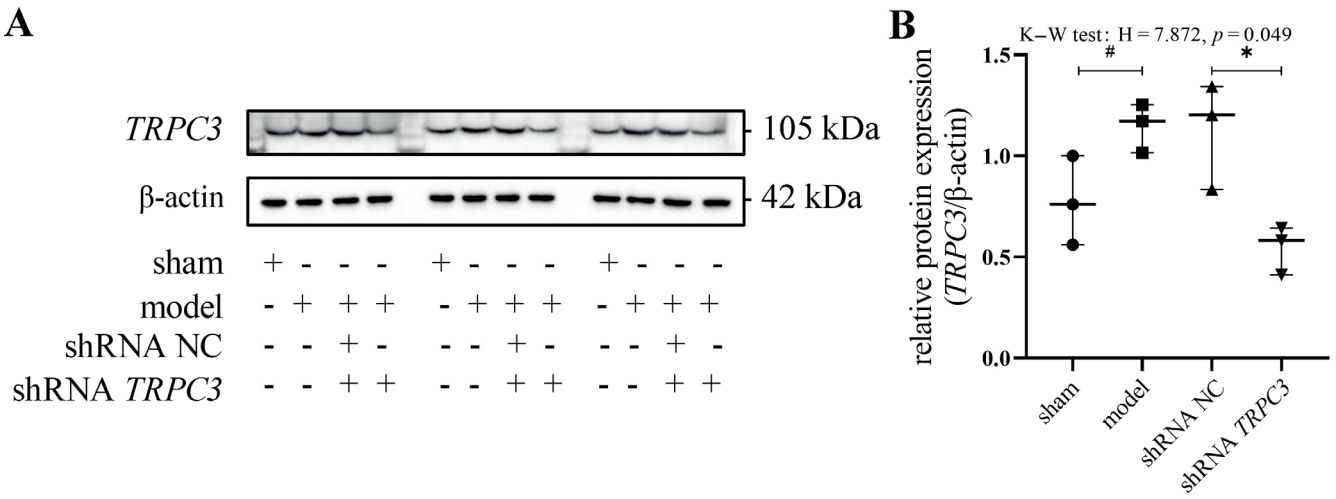

Fig. 3. Effect of *TRPC3* knockdown on *TRPC3* protein expression. A. *TRPC3* protein levels in intervertebral disc tissues were detected with western blot analysis; B. Quantitative analysis of *TRPC3* protein expression. Data are presented as the median (min–max) (n = 3). Data were analyzed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction

#p < 0.05 vs the sham group; *p < 0.05 vs the shRNA NC group; *TRPC3* – transient receptor potential canonical 3.

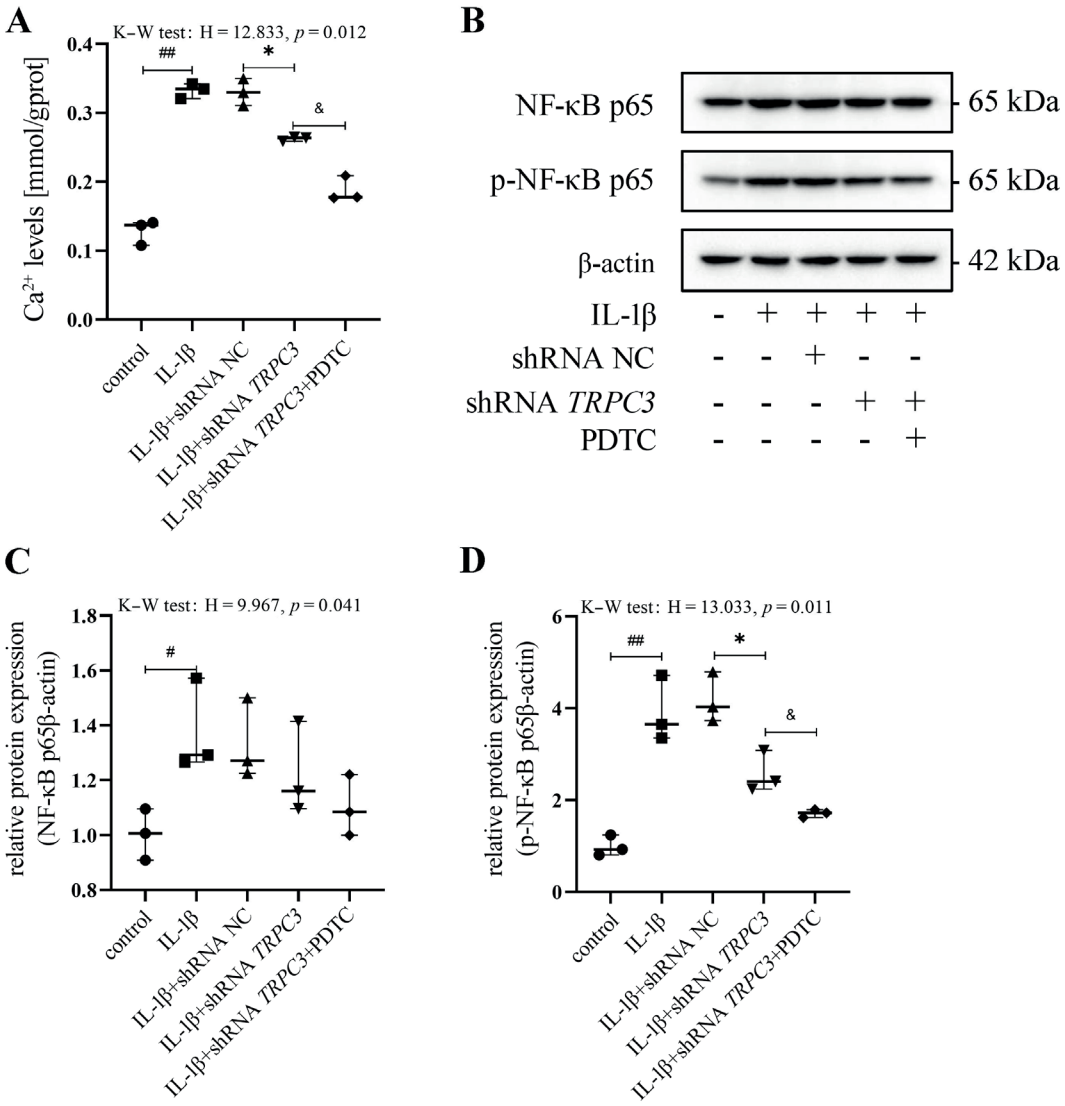

Fig. 4. *TRPC3* knockdown inhibits the Ca^{2+}/NF-κB pathway and promotes autophagy in NPCs. NPCs were pretreated with IL-1β, siRNA *TRPC3*, or the NF-κB inhibitor PDTC, respectively. A. Intracellular Ca^{2+} levels in NPCs were measured using a spectrophotometer; B. Western blot images showing NF-κB p65 and phosphorylated NF-κB p65 (p-NF-κB p65) protein expression in NPCs; C,D. Quantitative analysis of NF-κB p65 and p-NF-κB p65 protein expression. Data are presented as the median (minimum–maximum) (n = 3). Data were analyzed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction

##p < 0.01 vs the control group; *p < 0.05 vs the IL-1β + siRNA NC group; &p < 0.01 vs the IL-1β + siRNA *TRPC3* group; *TRPC3* – transient receptor potential canonical 3; NF-κB – nuclear factor kappa B; IL-1β – interleukin-1 beta; NPCs – nucleus pulposus cells; PDTC – pyrrolidine-dithiocarbamic acid.

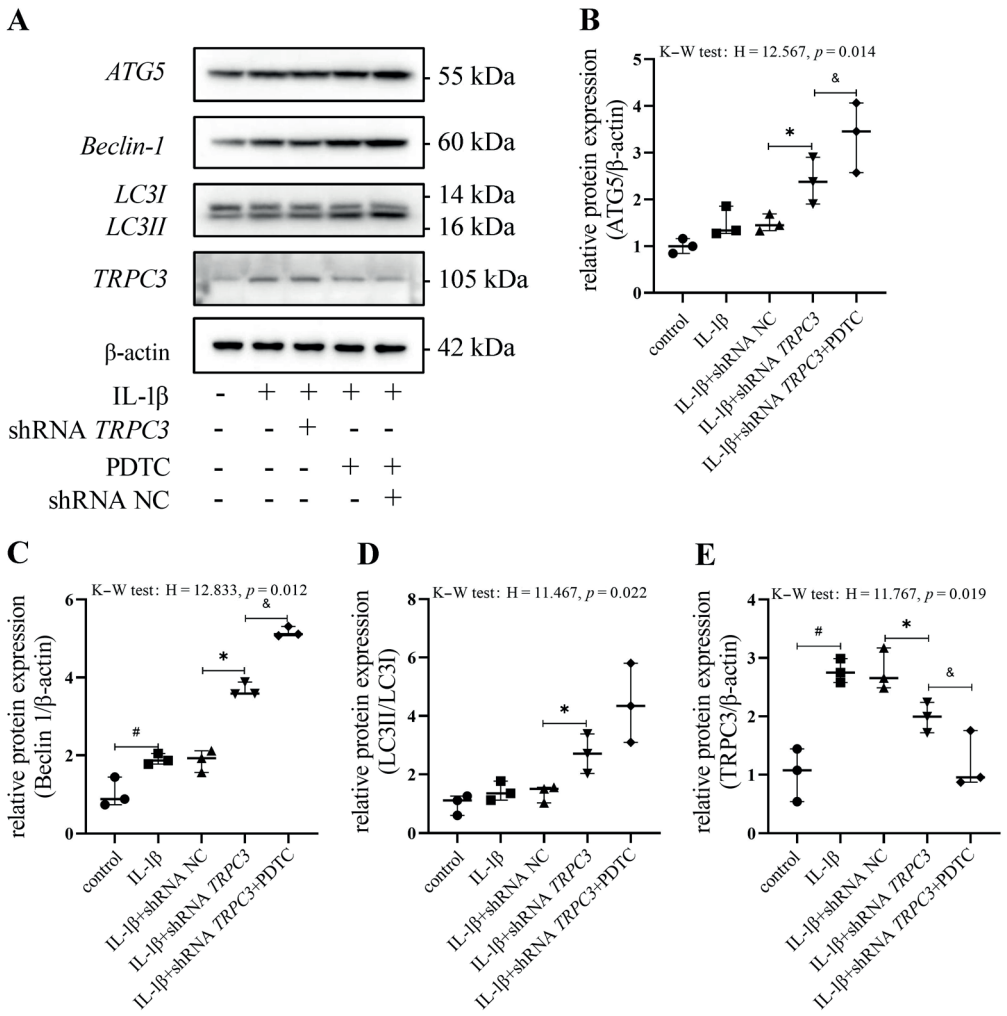

Fig. 5. Western blot analysis of *ATG5*, *Beclin-1*, *LC3-II/LC3-I*, and *TRPC3* expression in NPCs pretreated with *TRPC3* knockdown and the NF-κB inhibitor PDTC. A Representative western blot images of *ATG5*, *Beclin-1*, *LC3-II/LC3-I*, and *TRPC3* proteins; B–E Quantitative analysis of *ATG5*, *Beclin-1*, *LC3-II/LC3-I*, and *TRPC3* protein expression. Data are presented as the median (minimum–maximum) (n = 3). Data were analyzed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction

$^{\#}p < 0.05$ vs the control group; $^{*}p < 0.05$ vs. the IL-1β + shRNA NC group; $^{\&}p < 0.01$ vs the IL-1β + siRNA *TRPC3* group; *TRPC3* – transient receptor potential canonical 3; NF-κB – nuclear factor kappa B; *ATG5* – autophagy-related protein 5; NPCs – nucleus pulposus cells; PDTC – pyrrolidine thiocarbamate; IL-1β – interleukin-1 beta; NC – negative control.

The knockdown of *TRPC3* induced autophagy to regulate the expression of IDD-related proteins in NPCs

To investigate the regulatory effect of *TRPC3*-induced autophagy on ECM degradation in NPCs, western blot analysis was used to detect the expression levels of *MMP-13*, collagen II, and aggrecan (Fig. 8, Supplementary Table 6). Interleukin-1 beta treatment significantly reduced aggrecan (p = 0.007) and collagen II (p = 0.002) expression compared with the control group, while significantly increasing *MMP-13* expression (p = 0.002). In addition, knockdown of *TRPC3* reversed these expression changes (all p = 0.050). Compared with the IL-1β + siRNA *TRPC3* group, treatment with RAPA increased the expression of aggrecan and collagen II and decreased *MMP-13* expression (all p = 0.050). In contrast, treatment with 3-MA reduced collagen II expression (p = 0.050) and increased *MMP-13* expression (p = 0.050), while no significant difference was observed in aggrecan expression (p = 0.646). Taken together, these results indicate that *TRPC3* knockdown reverses IL-1β-induced ECM degradation by promoting autophagy in NPCs.

Discussion

Intervertebral disc degeneration is a leading cause of LBP and neurological compression syndromes. The pathogenesis of IDD is complex and mainly involves excessive mechanical stress, increased apoptosis of NPCs abnormal ECM degradation, dysregulated autophagy, oxidative stress-induced damage, and genetic factors. *TRPC3* is a member of the TRP family of cation channels, which are involved in various physiological processes, including calcium homeostasis and cell signaling. This study elucidates a novel role of *TRPC3* in IDD pathogenesis, demonstrating that *TRPC3* exacerbates disc degeneration by orchestrating Ca^{2+}/NF-κB-mediated suppression of autophagy, thereby providing a new research direction for targeted therapy of IDD.

The intervertebral disc is principally composed of 3 distinct components: the central gelatinous NP, the outer fibrous annulus fibrosus (AF), and the cartilaginous endplate (CEP).[35] The NP is predominantly composed of water, proteoglycans, and collagen, along with substantial amounts of elastic proteins, fibronectin, and laminin. In contrast, the AF is primarily composed of collagen fibers

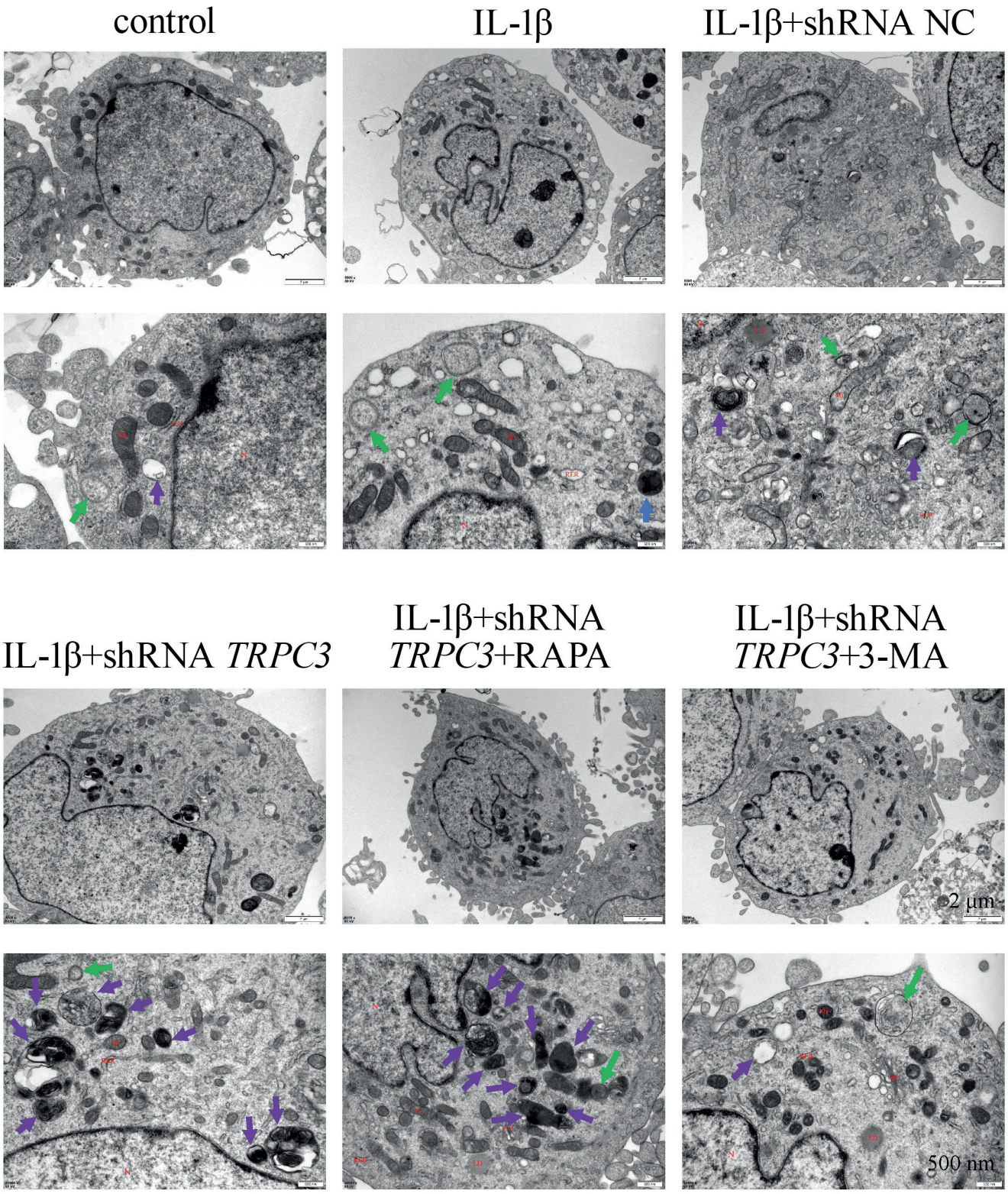

Fig. 6. Effect of *TRPC3* knockdown on autophagy in NPCs treated with IL-1β, siRNA *TRPC3*, the autophagy activator rapamycin (RAPA), and the autophagy inhibitor 3-methyladenine (3-MA). Transmission electron microscopy (TEM) images show that *TRPC3* knockdown enhances the formation of autophagosomes and primary lysosomes. Green arrows indicate autophagosomes, and purple arrows indicate autolysosomes. Scale bars = 2 μm and 500 nm

TRPC3 – transient receptor potential canonical 3; NPCs – nucleus pulposus cells; NC – negative control; 3-MA – 3-methyl adenine.

rich in type I collagen, whereas the CEP consists of hyaline cartilage.[36,37] In animal models, intervertebral disc tissue in IDD rats exhibits significant structural damage, including a reduction in NPCs, disorganized arrangement of the AF, and abnormal deposition of cartilaginous matrix. Inhibition of *TRPC3* ameliorates these pathological

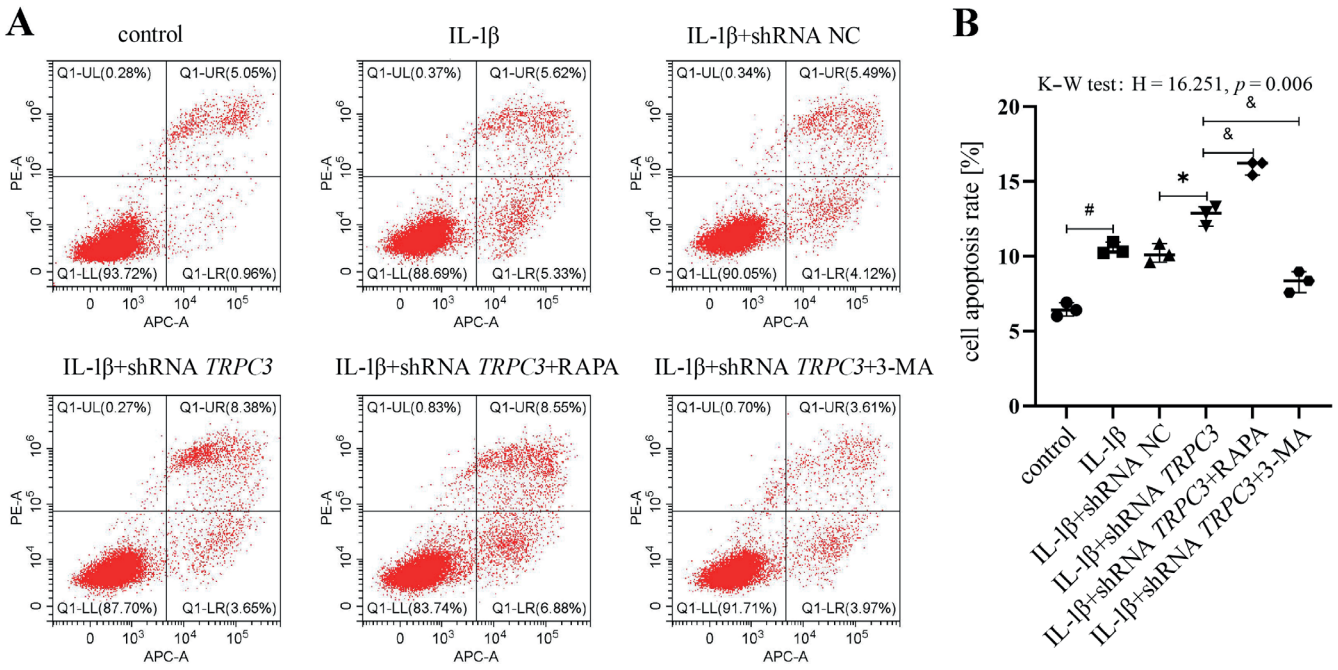

Fig. 7. Effect of *TRPC3* knockdown on apoptosis in NPCs. Cells were transiently transfected with siRNA negative control (NC) or siRNA *TRPC3* and then treated with IL-1β, the autophagy activator rapamycin (RAPA), or the autophagy inhibitor 3-methyladenine (3-MA), respectively. Apoptosis was analyzed with flow cytometry. Representative flow cytometry plots and quantitative results are shown. The Q1-UR region represents late apoptotic cells, Q1-UL represents necrotic cells, Q1-LL represents viable cells, and Q1-LR represents early apoptotic cells. Data are presented as the median (minimum–maximum) (n = 3). Data were analyzed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction

$^{\#}p < 0.05$ vs. the control group; $^{*}p < 0.05$ vs the IL-1β + shRNA NC group; $^{\&}p < 0.05$ vs the IL-1β + siRNA *TRPC3* group; *TRPC3* – transient receptor potential canonical 3; NPCs – nucleus pulposus cells; IL-1β – interleukin-1 beta.

changes in rat intervertebral disc tissue. These findings suggest that *TRPC3* primarily influences the progression of IDD by regulating cellular functions, such as autophagy and apoptosis, rather than by directly affecting bone structure. Consistent with this, previous studies have shown that members of the TRP channel family, such as TRPV4, regulate ECM homeostasis through ion channel activity in chondrocyte metabolism rather than directly modulating bone remodeling, thereby supporting the conclusions of the present study.[38]

Autophagy is a crucial cellular process that maintains homeostasis by degrading damaged components through lysosomes and plays a dual role in IDD: moderate autophagy eliminates damaged organelles and protects NPCs, whereas excessive autophagy may induce apoptosis.[19] ATG5 is a key regulator of autophagy, facilitating autophagosome formation and promoting autophagic flux.[39,40] The initiation of autophagy largely depends on *Beclin-1*, and *LC3-II* is commonly used as a marker of autophagy.[41–43] Our study reveals that under IL-1β stimulation, overexpression of *TRPC3* leads to a compensatory upregulation of the autophagy markers *ATG5*, *Beclin-1*, and *LC3-II*, accompanied by ECM degradation, as evidenced by downregulation of collagen II and aggrecan and upregulation of *MMP-13*. These findings suggest that *TRPC3* may cause abnormal accumulation of autophagosomes by blocking autophagic flux, possibly due to impaired lysosomal function. Further experiments confirmed that *TRPC3* knockdown or treatment with the autophagy inducer RAPA significantly restored autophagic activity and reversed ECM degradation, whereas the autophagy inhibitor 3-MA exacerbated matrix damage. Together, these results indicate that *TRPC3* promotes ECM metabolic imbalance by inhibiting protective autophagy, and that restoration of autophagic flux may be critical for delaying IDD.

Mechanistically, *TRPC3* acts as a nonselective cation channel, and its activation leads to intracellular Ca^{2+} overload.[44] Our data demonstrate that IL-1β stimulation significantly elevates intracellular Ca^{2+} levels in human NPCs, an effect that is effectively mitigated by *TRPC3* knockdown. This Ca^{2+} overload correlates with enhanced phosphorylation of NF-κB p65, suggesting that *TRPC3*-mediated Ca^{2+} influx is a critical activator of NF-κB signaling. In addition, Ca^{2+} may activate the IκB kinase (IKK) complex via calcium/calmodulin-dependent protein kinase II (CaMKII), thereby promoting IκBα degradation and subsequent nuclear translocation of NF-κB.[45] Notably, the NF-κB inhibitor PDTC not only suppressed p-NF-κB p65 phosphorylation but also attenuated IL-1β-induced *TRPC3* upregulation, revealing a self-reinforcing "*TRPC3*–Ca^{2+}–NF-κB–*TRPC3*" positive feedback loop. This vicious cycle amplifies apoptotic responses and ECM degradation, thereby promoting IDD progression. Knockdown of *TRPC3* or treatment with the NF-κB inhibitor PDTC reversed this phenomenon and synergistically enhanced the expression of *ATG5*, *Beclin-1*,

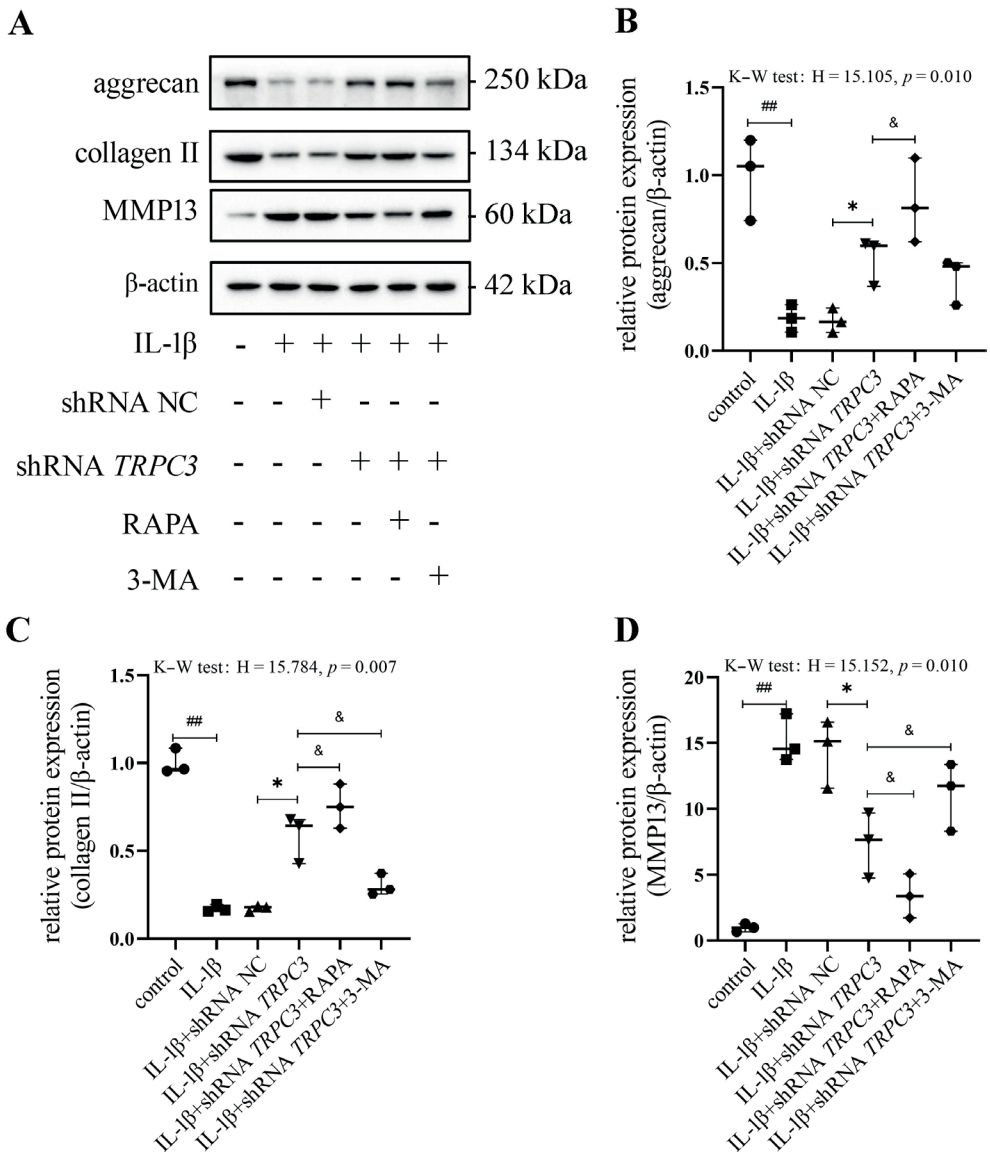

Fig. 8. *TRPC3* knockdown induces autophagy to regulate the expression of IDD-related proteins in NPCs. A Representative western blot images of *MMP-13*, collagen II, and aggrecan proteins; B–D Quantitative analysis of *MMP-13*, collagen II, and aggrecan protein expression. Data are presented as the median (minimum–maximum) (n = 3). Data were analyzed using the Kruskal–Wallis test followed by Dunn's post hoc test with Bonferroni correction

$^{\#\#}p < 0.01$ vs the control group; $^{*}p < 0.05$ vs the IL-1β + shRNA negative control (NC) group; $^{\&}p < 0.05$ vs the IL-1β + siRNA *TRPC3* group; *TRPC3* – transient receptor potential canonical 3; NPCs – nucleus pulposus cells; IDD – intervertebral disc degeneration; *MMP-13* – matrix metalloproteinase-13; IL-1β – interleukin-1 beta.

and *LC3-II*. These findings suggest that *TRPC3* negatively regulates autophagy through the Ca^{2+}/NF-κB axis, and that the underlying mechanism may involve NF-κB-mediated transcriptional repression of autophagy-related genes, such as the *ATG5* promoter.[46]

One of the characteristic features of IDD is apoptosis of NPCs.[47] The interplay between autophagy and apoptosis has attracted considerable attention in IDD research. In this study, we observed that IL-1β-induced apoptosis in NPCs was further exacerbated by *TRPC3* knockdown. In addition, the autophagy activator RAPA enhanced apoptosis, whereas the autophagy inhibitor 3-MA attenuated apoptotic responses. These findings suggest bidirectional regulation of autophagy at different stages: moderate autophagy inhibits apoptosis by clearing damaged mitochondria, whereas excessive autophagy – such as "autophagic stress" caused by impaired autophagic flux – may promote apoptosis by activating caspase pathways or depleting pro-survival signals.[48] *TRPC3* knockdown may trigger this latter effect by inducing excessive autophagy, including accumulation of autophagosomes.

Extracellular matrix degradation is a central feature of IDD,[49] primarily driven by an imbalance among MMPs, aggrecan, and collagen II. During the early stages of IDD, there is a notable increase in the production of type II collagen in the NP, suggesting an attempt by the tissue to initiate repair mechanisms. As degeneration progresses, type II collagen content markedly decreases, while type I collagen forms more pronounced fibrous structures in the AF and NP.[50] In addition, a key hallmark of degeneration is alteration of proteoglycans, particularly degradation of aggrecan, a major polymeric proteoglycan of the intervertebral disc.[51] Aggrecan, with its glycosaminoglycan side chains containing numerous negatively charged groups, confers unique permeability to the NP, allowing it to maintain a swollen state under compressive loads.[52] Moreover, recent studies have demonstrated that ECM degradation is an early and critical event in IDD,

involving metalloproteinases such as *MMP-1*, *MMP-3*, and *MMP-13*, which play key roles and are considered risk genes associated with disc degeneration.[53,54] Our study demonstrates that *TRPC3* knockdown reverses IL-1β-induced upregulation of *MMP-13* and downregulation of collagen II and aggrecan by activating autophagy. Mechanistically, autophagy may maintain ECM homeostasis through several pathways: 1) degrading damaged MMP precursors[55]; 2) inhibiting NF-κB-mediated *MMP-13* transcription via the mTORC1 pathway[56]; and 3) enhancing the anabolic function of NPCs to promote secretion of ECM components.[57] Notably, the autophagy inhibitor 3-MA partially reverses the protective effects of *TRPC3* knockdown on the ECM, further supporting the central role of autophagy. In addition, NF-κB, as a key transcription factor regulating *MMP-13* expression, may directly reduce ECM degradation when inhibited, while *TRPC3* knockdown exerts a synergistic protective effect through a dual mechanism involving autophagy activation and NF-κB inhibition.

Limitations of the study

This study has several limitations. First, we only conducted an initial exploration of the role of *TRPC3* in IDD. In addition, the rat IDD model cannot fully recapitulate the chronic degenerative process observed in humans; therefore, the effects and underlying mechanisms of *TRPC3* on ECM metabolism in IDD in vivo require further investigation. Although this study elucidates the mechanism by which *TRPC3* regulates autophagy through the Ca^{2+}/NF-κB pathway, the pathophysiology of IDD is complex and multifactorial, potentially involving multiple signaling pathways and molecular mechanisms beyond those examined here. Future research should explore additional signaling pathways and molecular mechanisms involved in IDD and their relationships with *TRPC3*. In addition, the potential involvement of other Ca^{2+} channels, such as *TRPV4*, cannot be excluded and warrants further validation of the functional specificity of *TRPC3* using gene-editing approaches, such as CRISPR–Cas9.

Conclusions

The results of this study indicate that *TRPC3* promotes autophagy by regulating the Ca^{2+}/NF-κB pathway, induces apoptosis of NPCs and ECM degradation, and thereby contributes to the development of IDD. These findings highlight the importance of *TRPC3* as a key therapeutic target in IDD and provide potential strategies for the treatment of this condition.

Future studies may focus on the development of small-molecule drugs targeting TRPC3, particularly highly selective TRPC3 inhibitors, as well as the application of nanodelivery technologies to enhance the clinical translation potential of TRPC3-targeted therapies for IDD. In addition, exploring combination strategies involving TRPC3 inhibitors together with NF-κB antagonists or autophagy modulators may synergistically enhance the suppression of IDD progression.

Supplementary files

The supplementary materials are available at https://doi.org/10.5281/zenodo.16352711. The package contains the following files:

Supplementary Table 1. Kruskal–Wallis test results for Fig. 2.

Supplementary Table 2. Kruskal–Wallis test results for Fig. 3.

Supplementary Table 3. Kruskal–Wallis test results for Fig. 4.

Supplementary Table 4. Kruskal–Wallis test results for Fig. 5.

Supplementary Table 5. Kruskal–Wallis test results for Fig. 7.

Supplementary Table 6. Kruskal–Wallis test results for Fig. 8.

Data Availability Statement

The raw data of the current study are openly available in Zenodo repository at https://doi.org/10.5281/zenodo.16352556.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Yingchao Gao: https://orcid.org/0009-0001-4869-9952
Ning Zhang: https://orcid.org/0009-0009-5142-9142
Jun-Fei Zhang: https://orcid.org/0000-0003-0575-4955
Zhengqi Fei: https://orcid.org/0009-0000-1266-3555

References

1. Xin J, Wang Y, Zheng Z, Wang S, Na S, Zhang S. Treatment of intervertebral disc degeneration. *Orthop Surg*. 2022;14(7):1271–1280. doi:10.1111/os.13254
2. Kamali A, Ziadlou R, Lang G, et al. Small molecule-based treatment approaches for intervertebral disc degeneration: Current options and future directions. *Theranostics*. 2021;11(1):27–47. doi:10.7150/thno.48987
3. Diwan AD, Melrose J. Intervertebral disc degeneration and how it leads to low back pain. *JOR Spine*. 2023;6(1):e1231. doi:10.1002/jsp2.1231
4. Xia Q, Zhao Y, Dong H, et al. Progress in the study of molecular mechanisms of intervertebral disc degeneration. *Biomed Pharmacother*. 2024;174:116593. doi:10.1016/j.biopha.2024.116593

5. Shen Y, Jiang Y, Jiang R, et al. Intervertebral disc degeneration mediates the causal effect of genetically predicted diffuse idiopathic skeletal hyperostosis on spinal stenosis: Evidence from a Mendelian randomization study. *JOR Spine*. 2025;8(1):e70041. doi:10.1002/jsp2.70041
6. Wang Y, Cheng H, Wang T, Zhang K, Zhang Y, Kang X. Oxidative stress in intervertebral disc degeneration: Molecular mechanisms, pathogenesis and treatment. *Cell Prolif*. 2023;56(9):e13448. doi:10.1111/cpr.13448
7. Fatoye F, Gebrye T, Ryan CG, Useh U, Mbada C. Global and regional estimates of clinical and economic burden of low back pain in high-income countries: A systematic review and meta-analysis. *Front Public Health*. 2023;11:1098100. doi:10.3389/fpubh.2023.1098100
8. Liu P, Ren X, Zhang B, Guo S, Fu Q. Investigating the characteristics of mild intervertebral disc degeneration at various age stages using single-cell genomics. *Front Cell Dev Biol*. 2024;12:1409287. doi:10.3389/fcell.2024.1409287
9. Yan SM, Han BQ, Song C, Yan LM. Molecular mechanisms and treatment strategies for discogenic lumbar pain. *Immunol Res*. 2025;73(1):111. doi:10.1007/s12026-025-09667-w
10. Song C, Hu P, Peng R, Li F, Fang Z, Xu Y. Bioenergetic dysfunction in the pathogenesis of intervertebral disc degeneration. *Pharmacol Res*. 2024;202:107119. doi:10.1016/j.phrs.2024.107119
11. Yurube T, Ito M, Kakiuchi Y, Kuroda R, Kakutani K. Autophagy and mTOR signaling during intervertebral disc aging and degeneration. *JOR Spine*. 2020;3(1):e1082. doi:10.1002/jsp2.1082
12. Wang Z, Li X, Yu P, et al. Role of autophagy and pyroptosis in intervertebral disc degeneration. *J Inflamm Res*. 2024;17:91–100. doi:10.2147/JIR.S434896
13. Liang H, Luo R, Li G, Zhang W, Song Y, Yang C. The proteolysis of ECM in intervertebral disc degeneration. *Int J Mol Sci*. 2022;23(3):1715. doi:10.3390/ijms23031715
14. Bianchi S, Bernardi S, Mattei A, et al. Morphological and biological evaluations of human periodontal ligament fibroblasts in contact with different bovine bone grafts treated with low-temperature deproteinisation protocol. *Int J Mol Sci*. 2022;23(9):5273. doi:10.3390/ijms23095273
15. Lu X, Lin Z, Li L, et al. Exosome-loaded methacrylated silk fibroin hydrogel delays intervertebral disc degeneration by DKK2-mediated mitochondrial unfolded protein response. *Chem Eng J*. 2025;511:162191. doi:10.1016/j.cej.2025.162191
16. Liu Y, Levine B. Autosis and autophagic cell death: The dark side of autophagy. *Cell Death Differ*. 2015;22(3):367–376. doi:10.1038/cdd.2014.143
17. Wang W, Sun Y, Tang P, et al. CircTBCK protects against osteoarthritis by regulating extracellular matrix and autophagy. *Human Cell*. 2025;38(2):60. doi:10.1007/s13577-025-01186-y
18. Zhang TW, Li ZF, Dong J, Jiang LB. The circadian rhythm in intervertebral disc degeneration: An autophagy connection. *Exp Mol Med*. 2020;52(1):31–40. doi:10.1038/s12276-019-0372-6
19. Kritschil R, Scott M, Sowa G, Vo N. Role of autophagy in intervertebral disc degeneration. *J Cell Physiol*. 2022;237(2):1266–1284. doi:10.1002/jcp.30631
20. Cheng ZR, Gan WK, Xiang Q, et al. Impaired degradation of PLCG1 by chaperone-mediated autophagy promotes cellular senescence and intervertebral disc degeneration. *Autophagy*. 2025;21(2):352–373. doi:10.1080/15548627.2024.2395797
21. Gruber HE, Hoelscher GL, Ingram JA, Bethea S, Hanley EN. Autophagy in the degenerating human intervertebral disc: In vivo molecular and morphological evidence, and induction of autophagy in cultured annulus cells exposed to proinflammatory cytokines. Implications for disc degeneration. *Spine (Phila Pa 1976)*. 2015;40(11):773–782. doi:10.1097/BRS.0000000000000865
22. Xie C, Shi Y, Chen Z, et al. Apigenin alleviates intervertebral disc degeneration via restoring autophagy flux in nucleus pulposus cells. *Front Cell Dev Biol*. 2022;9:787278. doi:10.3389/fcell.2021.787278
23. Zhang SJ, Yang W, Wang C, et al. Autophagy: A double-edged sword in intervertebral disk degeneration. *Clin Chim Acta*. 2016;457:27–35. doi:10.1016/j.cca.2016.03.016
24. Casas J, Meana C, López-López JR, Balsinde J, Balboa MA. Lipin-1-derived diacylglycerol activates intracellular TRPC3 which is critical for inflammatory signaling. *Cell Mol Life Sci*. 2021;78(24):8243–8260. doi:10.1007/s00018-021-03999-0
25. Song T, Hao Q, Zheng YM, Liu QH, Wang YX. Inositol 1,4,5-trisphosphate activates TRPC3 channels to cause extracellular Ca^{2+} influx in airway smooth muscle cells. *Am J Physiol Lung Cell Mol Physiol*. 2015;309(12):L1455–L1466. doi:10.1152/ajplung.00148.2015
26. Klein S, Mentrup B, Timmen M, et al. Modulation of transient receptor potential channels 3 and 6 regulates osteoclast function with impact on trabecular bone loss. *Calcif Tissue Int*. 2020;106(6):655–664. doi:10.1007/s00223-020-00673-8
27. Kim JE, Kim SY, Lim SY, Kieff E, Song YJ. Role of Ca^{2+}/calmodulin-dependent kinase II–IRAK1 interaction in LMP1-induced NF-κB activation. *Mol Cell Biol*. 2014;34(3):325–334. doi:10.1128/MCB.00912-13
28. Cheng P, Wei H, Chen H, Wang Z, Mao P, Zhang H. DNMT3-mediated methylation of PPARγ promote intervertebral disc degeneration by regulating the NF-κB pathway. *J Cell Mol Med*. 2024;28(2):e18048. doi:10.1111/jcmm.18048
29. Chen F, Liu H, Wang X, et al. Melatonin activates autophagy via the NF-κB signaling pathway to prevent extracellular matrix degeneration in intervertebral disc. *Osteoarthritis Cartilage*. 2020;28(8):1121–1132. doi:10.1016/j.joca.2020.05.011
30. Sun Z, Tang X, Wang H, et al. LncRNA *H19* aggravates intervertebral disc degeneration by promoting the autophagy and apoptosis of nucleus pulposus cells through the miR-139/CXCR4/NF-κB axis. *Stem Cells Dev*. 2021;30(14):736–748. doi:10.1089/scd.2021.0009
31. Huang Y, Wang L, Luo B, et al. Associations of lumber disc degeneration with paraspinal muscles myosteatosis in discogenic low back pain. *Front Endocrinol (Lausanne)*. 2022;13:891088. doi:10.3389/fendo.2022.891088
32. Zou Z, Hu X, Luo T, et al. Naturally-occurring spinosyn A and its derivatives function as argininosuccinate synthase activator and tumor inhibitor. *Nat Commun*. 2021;12(1):2263. doi:10.1038/s41467-021-22235-8
33. Zou R, Shi W, Chang X, et al. The DNA-dependent protein kinase catalytic subunit exacerbates endotoxemia-induced myocardial microvascular injury by disrupting the MOTS-c/JNK pathway and inducing profilin-mediated lamellipodia degradation. *Theranostics*. 2024;14(4):1561–1582. doi:10.7150/thno.92650
34. Lee Y, DiMaulo-Milk E, Leslie J, Ding L. Hematopoietic stem cells temporally transition to thrombopoietin dependence in the fetal liver. *Sci Adv*. 2022;8(11):eabm7688. doi:10.1126/sciadv.abm7688
35. Sun Z, Liu B, Luo ZJ. The immune privilege of the intervertebral disc: Implications for intervertebral disc degeneration treatment. *Int J Med Sci*. 2020;17(5):685–692. doi:10.7150/ijms.42238
36. Marfia G, Campanella R, Navone SE, et al. Potential use of human adipose mesenchymal stromal cells for intervertebral disc regeneration: A preliminary study on biglycan-deficient murine model of chronic disc degeneration. *Arthritis Res Ther*. 2014;16(5):457. doi:10.1186/s13075-014-0457-5
37. Lakstins K, Arnold L, Gunsch G, et al. Characterization of the human intervertebral disc cartilage endplate at the molecular, cell, and tissue levels. *J Orthop Res*. 2021;39(9):1898–1907. doi:10.1002/jor.24854
38. O'Conor CJ, Leddy HA, Benefield HC, Liedtke WB, Guilak F. TRPV4-mediated mechanotransduction regulates the metabolic response of chondrocytes to dynamic loading. *Proc Natl Acad Sci U S A*. 2014;111(4):1316–1321. doi:10.1073/pnas.1319569111
39. Wang F, Trosdal ES, Paddar MA, et al. The role of ATG5 beyond Atg8ylation and autophagy. *Autophagy*. 2024;20(2):448–450. doi:10.1080/15548627.2023.2273703
40. Changotra H, Kaur S, Yadav SS, Gupta GL, Parkash J, Duseja A. ATG5: A central autophagy regulator implicated in various human diseases. *Cell Biochem Funct*. 2022;40(7):650–667. doi:10.1002/cbf.3740
41. Li G, Rao H, Xu W. Puerarin plays a protective role in chondrocytes by activating Beclin1-dependent autophagy. *Biosci Biotechnol Biochem*. 2021;85(3):621–625. doi:10.1093/bbb/zbaa078
42. Zhang Y, Cui Y, Wang L, Han J. Autophagy promotes osteoclast podosome disassembly and cell motility athrough the interaction of kindlin3 with LC3. *Cell Signal*. 2020;67:109505. doi:10.1016/j.cellsig.2019.109505
43. Quan M, Hong M, Ko M, Kim Y. Relationships between disc degeneration and autophagy expression in human nucleus pulposus. *Orthop Surg*. 2020;12(1):312–320. doi:10.1111/os.12573
44. Fliniaux I, Germain E, Farfariello V, Prevarskaya N. TRPs and Ca^{2+} in cell death and survival. *Cell Calcium*. 2018;69:4–18. doi:10.1016/j.ceca.2017.07.002

45. Li T, Meng Y, Ding P, et al. Pathological implication of CaMKII in NF-κB pathway and SASP during cardiomyocytes senescence. *Mech Ageing Dev*. 2023;209:111758. doi:10.1016/j.mad.2022.111758
46. Peng X, Wang Y, Li H, et al. ATG5-mediated autophagy suppresses NF-κB signaling to limit epithelial inflammatory response to kidney injury. *Cell Death Dis*. 2019;10(4):253. doi:10.1038/s41419-019-1483-7
47. Wang D, He X, Wang D, et al. Quercetin suppresses apoptosis and attenuates intervertebral disc degeneration via the SIRT1-autophagy pathway. *Front Cell Dev Biol*. 2020;8:613006. doi:10.3389/fcell.2020.613006
48. Lee DY, Bahar ME, Kim CW, et al. Autophagy in osteoarthritis: A double-edged sword in cartilage aging and mechanical stress response. A systematic review. *J Clin Med*. 2024;13(10):3005. doi:10.3390/jcm13103005
49. Kaneda G, Zila L, Wechsler JT, et al. What a pain in the back: Etiology, diagnosis and future treatment directions for discogenic low back pain. *Bone Res*. 2025;13(1):89. doi:10.1038/s41413-025-00472-7
50. Hua WB, Wu XH, Zhang YK, et al. Dysregulated miR-127-5p contributes to type II collagen degradation by targeting matrix metalloproteinase-13 in human intervertebral disc degeneration. *Biochimie*. 2017;139:74–80. doi:10.1016/j.biochi.2017.05.018
51. Gruber HE, Hoelscher GL, Ingram JA, Bethea S, Zinchenko N, Hanley EN. Variations in aggrecan localization and gene expression patterns characterize increasing stages of human intervertebral disk degeneration. *Exp Mol Pathol*. 2011;91(2):534–539. doi:10.1016/j.yexmp.2011.06.001
52. Sivan SS, Wachtel E, Roughley P. Structure, function, aging and turnover of aggrecan in the intervertebral disc. *Biochim Biophys Acta Gen Subj*. 2014;1840(10):3181–3189. doi:10.1016/j.bbagen.2014.07.013
53. Basaran R, Senol M, Ozkanli S, Efendioglu M, Kaner T. Correlation of matrix metalloproteinase (MMP)-1, -2, -3, and -9 expressions with demographic and radiological features in primary lumbar intervertebral disc disease. *J Clin Neurosci*. 2017;41:46–49. doi:10.1016/j.jocn.2017.03.001
54. Zou X, Zhang X, Han S, et al. Pathogenesis and therapeutic implications of matrix metalloproteinases in intervertebral disc degeneration: A comprehensive review. *Biochimie*. 2023;214:27–48. doi:10.1016/j.biochi.2023.05.015
55. Li WD, Li NP, Song DD, Rong JJ, Qian AM, Li XQ. Metformin inhibits endothelial progenitor cell migration by decreasing matrix metalloproteinases, MMP-2 and MMP-9, via the AMPK/mTOR/autophagy pathway. *Int J Mol Med*. 2017;39(5):1262–1268. doi:10.3892/ijmm.2017.2929
56. Huang P, Dong RY, Wang P, Xu M, Sun X, Dong XP. MCOLN/TRPML channels in the regulation of MTORC1 and autophagy. *Autophagy*. 2024;20(5):1203–1204. doi:10.1080/15548627.2023.2300922
57. Wang K, Yao D, Li Y, et al. TAK-715 alleviated IL-1β-induced apoptosis and ECM degradation in nucleus pulposus cells and attenuated intervertebral disc degeneration ex vivo and in vivo. *Arthritis Res Ther*. 2023;25(1):45. doi:10.1186/s13075-023-03028-4

Balancing osteogenesis and adipogenesis in osteogenesis imperfecta: *PLIN2* and *E2F2* as key targets for mesenchymal stem cell therapy

Xishun Wang[C,D,F], Zhenjiang Liu[A–C,E,F], Xinyong Hu[B,C,E,F], Yinpeng Cui[B,C,E,F]

Department of Orthopedics, Capital Center for Children's Health, Capital Medical University, 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(5):835–846

Address for correspondence
Xishun Wang
E-mail: wangxishun96@163.com

Funding sources
None declared

Conflict of interest
None declared

Received on April 27, 2025
Reviewed on July 7, 2025
Accepted on July 31, 2025

Published online on May 26, 2026

Abstract

Background. Osteogenesis imperfecta (OI) necessitates innovative mesenchymal stem cell (MSC) therapies targeting key molecular pathways to enhance targeted and combination treatments and improve bone health.

Objectives. To investigate the therapeutic mechanisms of various interventions for OI by analyzing relevant datasets, with a focus on lipid metabolism-related genes, particularly *PLIN2*, in order to determine whether they influence the balance between osteoblast and adipocyte differentiation.

Materials and methods. This study analyzed datasets from the Gene Expression Omnibus (GEO; GSE157587, GSE214064, GSE186141) and UK Biobank genome-wide association study (GWAS) summary statistics (UKB-b-4657, UKB-b-1096, UKB-b-8875, UKB-b-20124) using bioinformatics tools, including GEO2R, DESeq2, TwoSampleMR, MR-Egger, MR-PRESSO, gwasrapidd, and summary data-based Mendelian randomization (MR), to identify differentially expressed genes (DEGs) and assess causal relationships with heel bone mineral density (BMD).

Results. Differentially expressed genes analysis of GSE157587 identified *PLIN2* as the most significant gene influenced by MSC therapy in OI (log2 fold change = 0.428, adjusted $p = 3.29 \times 10^{-6}$), whereas GSE186141 revealed 770 DEGs in OI patients, with 7 overlapping with *PLIN2*-related genes. Notably, *TNFRSF19* (log2 fold change = –2.7454, adjusted $p = 3.930 \times 10^{-7}$ in OI; 1.5001, adjusted $p = 3.482 \times 10^{-3}$ in *PLIN2* knockdown) and *E2F2* (log2 fold change = –2.1428, adjusted $p = 8.830 \times 10^{-5}$ in OI; 1.7207, adjusted $p = 1.244 \times 10^{-2}$ in *PLIN2* knockdown) were identified as key genes. Mendelian randomization analysis confirmed a negative association between *E2F2* and heel BMD ($p = 1.116 \times 10^{-7}$ to 6.073×10^{-5}; effect size –0.0461 to –0.0277).

Conclusions. *PLIN2* and *E2F2* emerge as critical targets for refining MSC therapy in OI, with the potential to improve bone formation and reduce fat accumulation by restoring the osteogenesis–adipogenesis balance. These findings may support the development of combination therapies or engineered MSCs, ultimately improving clinical outcomes for patients with OI.

Key words: Mendelian randomization, lipid metabolism, osteogenesis imperfecta, mesenchymal stem cell therapy, bone homeostasis

Cite as
Wang X, Liu Z, Hu X, Cui Y. Balancing osteogenesis and adipogenesis in isteogenesis imperfecta: *PLIN2* and *E2F2* as key targets for mesenchymal stem cell therapy. *Adv Clin Exp Med*. 2026;35(5):835–846. doi:10.17219/acem/208840

DOI
10.17219/acem/208840

Highlights

- *PLIN2* modulates lipid metabolism and mesenchymal stem cell (MSC) differentiation in osteogenesis imperfecta.
- *E2F2* is a key regulator of bone cell function and bone mineral density.
- Integrated transcriptomic and Mendelian randomization analyses identify *PLIN2* and *E2F2* as therapeutic targets.
- MSC therapy in osteogenesis imperfecta remains limited by variable and transient clinical effects.

Background

Osteogenesis imperfecta (OI) is a genetic disorder characterized by bone fragility and impaired bone formation. This rare connective tissue disorder affects approx. 11.6 per 100,000 pediatric individuals, as reported in a recent nationwide registry study conducted in Turkey (data from 2016–2022).[1] Current treatments for OI primarily rely on bisphosphonates and other antiresorptive agents. Despite these approaches, significant challenges remain, including limited efficacy in severe cases, long-term adverse effects, and the inability to fully restore normal bone quality and reduce fracture risk. Therefore, the identification of novel therapeutic targets remains essential for improving the management of this serious disease. Recent studies have highlighted the complex interplay between glucose and lipid metabolism and bone homeostasis. Notably, hypoglycemic agents such as metformin[2] and glucagon-like peptide-1 (GLP-1) receptor agonists[3,4] have demonstrated potential pro-osteogenic effects, while lipid-lowering agents such as statins[5,6] have also shown promise in promoting bone formation. These findings suggest that targeting glucose and lipid metabolic pathways may provide novel therapeutic strategies for OI.

Recent studies have highlighted the complex interplay between glucose and lipid metabolism and bone homeostasis. Notably, high-fat diet-induced glucose intolerance leads to dysregulated osteoblast lipid metabolism, resulting in reduced bone formation; this effect can be mitigated by enhanced fatty acid oxidation.[7] Furthermore, extracellular metabolites such as lactate, as well as interactions with mesenchymal stromal cells in the tumor microenvironment (TME), exacerbate lipid droplet accumulation in osteosarcoma cells. Targeting PLIN2, a key protein associated with lipid droplet formation, significantly reduces cell viability and increases reactive oxygen species (ROS) production in these cells.[8] PLIN2, a protein involved in lipid droplet formation and fatty acid oxidation, has emerged as a critical regulator in lipid metabolism. Specifically, *PLIN2* knockdown has been shown to induce lipolysis and increase fatty acid oxidation, making it a promising target in various metabolic disorders.[9,10] Given the dynamic relationship between adipogenesis and osteogenesis, and the potential role of *PLIN2* in regulating lipid metabolism and bone cell differentiation, we hypothesize that therapeutic interventions for OI may exert their effects, at least in part, by modulating *PLIN2* expression and thereby influencing the balance between osteogenesis and adipogenesis.

Objectives

This study aims to investigate the therapeutic mechanisms of various interventions for OI by analyzing relevant datasets, with a focus on lipid metabolism-related genes, particularly *PLIN2*, to determine whether they influence the balance between osteoblast and adipocyte differentiation. By elucidating the role of *PLIN2* in OI pathogenesis and therapeutic responses, this study seeks to provide novel insights into potential therapeutic targets for OI.

Materials and methods

Datasets

We retrieved datasets from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) database to investigate innovative therapeutic strategies aimed at improving bone health in patients with OI. Among these, mesenchymal stem cell (MSC) therapy has demonstrated potential in clinical settings, with studies showing improvements in bone parameters and stimulation of pro-osteogenic responses (GSE157587), whereas other emerging therapies remain largely limited to animal or cell line models. The dataset GSE214064 examines the impact of *PLIN2* deficiency on hepatic lipid droplet storage and gene expression in fasted mice, demonstrating that *PLIN2* plays a crucial role in regulating lipid droplet size and neutral lipid storage under fasting conditions. The dataset GSE186141 investigates mutations in the *COL1A1* and *COL1A2* genes and their correlation with clinical phenotypes in a Chinese cohort of patients with OI.

The included datasets comprise genome-wide association study (GWAS) summary statistics related to heel bone mineral densite (BMD), derived from the UK Biobank, a large-scale prospective cohort study. Specifically, UKB-b-4657 was designated as the discovery dataset, whereas UKB-b-1096, UKB-b-8875, and UKB-b-20124 served as validation datasets. All datasets originate from a European population, including both males and females, and were generated in 2018.

The variables are continuous and expressed in standard deviations (SD). Sample sizes vary across datasets, ranging from approx. 146,000 to 265,000 individuals, and each dataset includes nearly 9.85 million single-nucleotide polymorphisms (SNPs). These datasets were generated using PHESANT-derived variables from the UK Biobank, based on the hg19/GRCh37 reference genome build. Notably, UKB-b-20124 represents heel BMD T-scores, whereas the remaining datasets correspond to raw BMD measurements.

DEGs analysis methodology

To address the hypothesis that MSC therapy or *PLIN2* deficiency alters gene expression profiles relevant to OI pathophysiology, we employed a dual approach using both the online tool GEO2R (https://www.ncbi.nlm.nih.gov/geo/geo2r) and the R package DESeq2 (v. 3.21.1; https://github.com/thelovelab/DESeq2).[11] DESeq2 models count data using a negative binomial generalized linear model, which accounts for the variance–mean dependence inherent in RNA-seq data and does not assume normality of raw counts. All DESeq2 analyses were performed using standard workflows and default parameters unless otherwise specified. The primary criteria for identifying differentially expressed genes (DEGs) were defined as a log2 fold change >1 or <–1 and an adjusted $p < 0.05$. For multiple testing correction, DESeq2 applies the Benjamini–Hochberg procedure to control the false discovery rate (FDR). However, when these stringent log2 fold change thresholds did not yield any DEGs, the selection criteria were relaxed to include genes with an adjusted $p < 0.05$. This approach ensured the inclusion of potentially relevant genes exhibiting statistically significant differential expression, even with more modest effect sizes.

Data acquisition and preparation

To investigate potential causal relationships between overlapping DEGs identified from OI and *PLIN2*-related gene expression profiles and heel BMD, we performed Mendelian randomization (MR) analysis. We first identified overlapping DEGs from the GSE186141 dataset, which contains gene expression data from patients with OI. These overlapping genes were subsequently analyzed for their association with heel BMD. GWAS summary statistics for heel BMD were obtained from the UK Biobank. One dataset was designated as the discovery dataset, whereas the remaining datasets served as validation datasets. To facilitate efficient retrieval and management of GWAS data, we used the gwasrapidd R package,[12] which enables querying and downloading data from the GWAS Catalog.

MR analysis

To address our hypothesis regarding the causal effects of gene expression on heel BMD, we performed two-sample MR analysis using the TwoSampleMR R package (v. 0.6.17; R Foundation for Statistical Computing, Vienna, Austria).[13] This approach uses genetic variants as instrumental variables to infer causal relationships between exposures (gene expression) and outcomes (heel BMD). For each gene, independent SNPs associated with gene expression were selected as instrumental variables. Instrumental variables were selected based on stringent criteria to ensure validity: SNPs were required to have a $p < 5 \times 10^{-8}$ for association with gene expression and an F-statistic >10 to minimize weak instrument bias. To account for linkage disequilibrium (LD), clumping was performed using PLINK (v. 1.9; https://www.cog-genomics.org/plink) with a 1000 Genomes Project European reference panel, applying an LD r^2 threshold of 0.3 and a clumping window of 1,000 kb to ensure independence of selected SNPs. This procedure reduces multicollinearity among instrumental variables by retaining only independent variants. The inverse variance weighted (IVW) method was used as the primary approach to estimate causal effects, as it provides the most precise estimates when all instrumental variable assumptions are satisfied. This method performs a weighted linear regression of SNP–outcome associations on SNP–exposure associations, with weights inversely proportional to the variance of the SNP–outcome effects. The validity of MR analysis relies on 3 core assumptions for the instrumental variables: 1) relevance (SNPs are associated with the exposure), ensured by the applied p-value and F-statistic thresholds; 2) independence (SNPs are not associated with confounders of the exposure–outcome relationship), supported by the random allocation of genetic variants and further evaluated using MR-Egger regression to detect directional pleiotropy; and 3) exclusion restriction (SNPs influence the outcome only through the exposure), assessed using heterogeneity tests. Regarding the underlying statistical framework, MR assumes linear relationships, homoscedasticity (constant variance of residuals), and normally distributed residuals for valid inference of standard errors (SEs) and p-values. Although these assumptions cannot be directly verified using summary-level data, the application of complementary robust MR methods (e.g., MR-Egger and MR-PRESSO (Mendelian Randomization Pleiotropy RESidual Sum and Outlier)), along with validation across multiple datasets, enhances the robustness of the findings to potential violations. All analyses using the TwoSampleMR package were conducted with default parameters unless otherwise specified.

Statistical analyses and heterogeneity assessment

Causal effects of gene expression on heel BMD were estimated using the IVW method, and statistical significance was assessed using p-values. Heterogeneity among instrumental variables was evaluated using Cochran's Q statistic. Directional pleiotropy was assessed

using MR-Egger regression. Additionally, MR-PRESSO analysis was performed to detect and correct for horizontal pleiotropy.[14] A key consideration in MR analyses based on summary-level data is the potential for bias due to sample overlap between exposure and outcome GWAS datasets or residual linkage disequilibrium if clumping is insufficient. Although strict clumping parameters were applied and findings were validated across multiple UK Biobank datasets to mitigate these concerns, these limitations are inherent to analyses based on summary-level data.

SMR integration

To further validate our findings and integrate GWAS summary statistics with expression quantitative trait loci (eQTL) data, we applied the summary data-based Mendelian randomization (SMR) method using default parameters for heterogeneity ($I^2 < 0.05$ and P_HEIDI > 0.01).[15] SMR enables the identification of candidate genes for complex traits by integrating GWAS summary statistics with eQTL data. Specifically, SMR tests for pleiotropy, whereby a genetic variant influences both gene expression and a complex trait through a shared underlying causal variant. By applying multiple MR approaches and integrating GWAS and eQTL data, we aimed to comprehensively assess potential causal relationships between the identified overlapping DEGs and heel BMD.

Results

DEGs of MSC therapy on OI

Differentially expressed gene analysis was performed using GEO2R based on the GSE157587 dataset, which examines the effects of MSC therapy in OI. This analysis identified a limited set of 4 potentially significant DEGs: *PLIN2*, *PDK4*, *ANGPTL4*, and *HADHA*. Notably, the log2 fold change values for these genes were modest, indicating moderate changes in gene expression. Among the identified DEGs, *PLIN2* showed the most significant adjusted p-value (adjusted $p = 3.290 \times 10^{-6}$), with a corresponding p-value of 1.920×10^{-10}, lfcSE of 0.0673, test statistic of 6.367, log2 fold change of 0.428, and a baseMean of 6529.790 (Table 1).

DEGs analysis of *PLIN2*-deficient mice reveals significant transcriptomic alterations

Differentially expressed gene analysis was conducted on the GSE214064 dataset, which examined gene expression differences between $PLIN2^{+/+}$ and $PLIN2^{-/-}$ mice under 24-h fasting conditions in a *C57BL/6N* background. The analysis aimed to identify *PLIN2*-related genes. Transcriptomic data clearly distinguished the $PLIN2^{+/+}$ and $PLIN2^{-/-}$ sample groups from each other (Fig. 1A). The results revealed 43 downregulated and 102 upregulated

Table 1. Differentially expressed genes (DEGs) following MSC therapy in OI

Gene symbol	GeneID	p_{adj}	p-value	lfcSE	stat	log2FoldChange	baseMean	Description
PLIN2	123	0.001	1.92E-10	0.067	6.367	4.28E-01	6529.790	*PLIN2* (perilipin 2): Perilipin 2, involved in lipid droplet formation and storage, may influence MSC differentiation and lipid metabolism, potentially affecting bone matrix quality and osteoblast function OI following MSC therapy.
PDK4	5166	0.001	4.46E-08	0.164	5.472	8.98E-01	2173.820	PDK4 (pyruvate dehydrogenase kinase 4): Pyruvate dehydrogenase kinase 4, a regulator of glucose and energy metabolism, can modulate the metabolic environment of bone-forming cells, possibly enhancing MSC-mediated repair of defective bone tissue in OI.
ANGPTL4	51129	0.010	1.77E-06	0.142	4.778	6.79E-01	6528.740	ANGPTL4 (angiopoietin like 4): Angiopoietin like 4, known for its role in angiogenesis and lipid metabolism, might contribute to improved vascular support and tissue remodeling in OI, potentially aiding MSC therapy in promoting bone regeneration.
HADHA	3030	0.014	3.20E-06	0.033	4.658	1.51E-01	7701.840	HADHA (hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit alpha): Hydroxyacyl-CoA dehydrogenase trifunctional multienzyme complex subunit alpha, essential for fatty acid oxidation, may support the energy demands of MSCs during osteogenic differentiation, potentially improving bone strength in OI post-therapy.

Gene symbol – official gene symbol; GeneID – NCBI gene identifier; p_{adj} – adjusted p-value (Benjamini–Hochberg FDR); p-value – raw p-value (Wald test); lfcSE – log2 fold change standard error; stat – Wald statistic; log2FoldChange – log2 fold change (MSC therapy vs control); baseMean – mean of normalized counts across all samples; description – gene name and functional annotation; MSC – mesenchymal stem cells; OI – osteogenesis imperfecta; NCBI – National Center for Biotechnology Information.

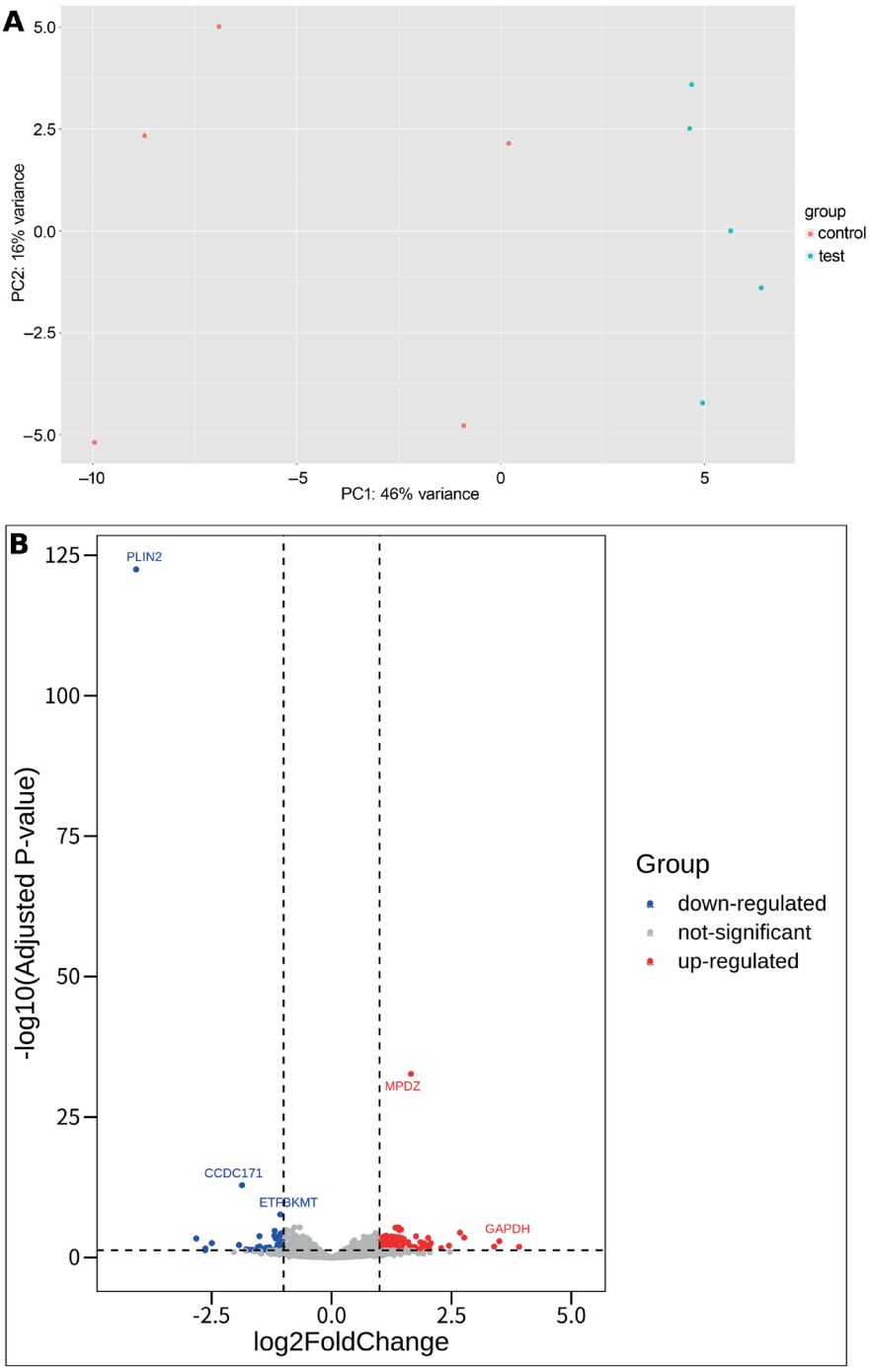

Fig. 1. Differentially expressed genes (DEGs) analysis in *PLIN2*-deficient mice. A. Principal component analysis (PCA) plot of *PLIN2*$^{+/+}$ and *PLIN2*$^{-/-}$ samples based on variance-stabilizing transformation, demonstrating clear separation between groups; B. Volcano plot of DEGs between *PLIN2*$^{+/+}$ and *PLIN2*$^{-/-}$ mice, highlighting significantly upregulated and downregulated genes

DEGs with statistical significance. Among the downregulated genes, *PLIN2* itself exhibited the most substantial log2 fold change, with a value of −4.074 and an adjusted $p = 3.180 \times 10^{-123}$. Other notable downregulated genes included *CCDC171*, *ETFBKMT*, and *CCNG2*, each demonstrating log2 fold changes greater than −1 and adjusted $p < 0.001$. Conversely, the upregulated gene list comprised 102 genes, with *MPDZ* showing the highest log2 fold change of 1.655 and an adjusted $p = 2.100 \times 10^{-33}$. Other significantly upregulated genes included *S100A9*, *ZBP1*, and *CIRBP*, all with log2 fold changes exceeding 1 and adjusted $p < 0.001$ (Fig. 1B).

Overlap of DEGs between OI and *PLIN2*-related gene expression profiles

Differentially expressed gene analysis was performed on the GSE214064 dataset, which examined gene expression differences between *PLIN2*$^{+/+}$ and *PLIN2*$^{-/-}$ mice under 24-h fasting conditions in a *C57BL/6N* background. The analysis aimed to identify *PLIN2*-related genes. Transcriptomic profiling clearly distinguished *PLIN2*$^{+/+}$ and *PLIN2*$^{-/-}$ samples (Fig. 1A). A total of 43 downregulated and 102 upregulated DEGs were identified as statistically significant. Among the downregulated genes, *PLIN2*

exhibited the most substantial log2 fold change (–4.076; adjusted p = 3.180×10^{-123}). Other notable downregulated genes included *CCDC171*, *ETFBKMT*, and *CCNG2*, each with log2 fold change <–1 and adjusted p < 0.001. Conversely, 102 genes were significantly upregulated, with *MPDZ* showing the highest log2 fold change (1.655; adjusted p = 2.100×10^{-33}). Other significantly upregulated genes included *S100A9*, *ZBP1*, and *CIRBP*, all with log2 fold change >1 and adjusted p < 0.001 (Fig. 1B). The overlap between the *PLIN2*-related DEGs and genes previously associated with OI is presented in Table 2.

MR analysis results for overlap of DEGs and heel BMD

Mendelian randomization analysis of *E2F2* was conducted to evaluate its potential causal relationship with heel BMD using UK Biobank data. The UKB-b-4657 dataset served as the discovery dataset, whereas UKB-b-8875, UKB-b-20124, and UKB-b-1096 were used as validation datasets. The IVW method was applied. For each dataset, 29 SNPs were used as instrumental variables. The estimated effect sizes (β coefficients) consistently indicated a negative association between *E2F2* expression and heel BMD, suggesting that increased *E2F2* expression is associated with lower heel BMD. The p-values across all 4 datasets were highly significant, ranging from 1.116×10^{-7} to 6.073×10^{-5}, indicating strong statistical evidence for this association. The 95% confidence intervals (95% CIs) for the effect estimates further supported this finding, as all intervals indicated negative effects across datasets. The corresponding odds ratios (ORs) were close to 1 (0.955–0.973), suggesting that although the association is statistically significant, the magnitude of the effect is relatively small (Table 3, Fig. 2). No other DEGs showed a causal relationship with heel BMD using the IVW method.

Mendelian randomization analysis of the causal relationship between *E2F2* and heel BMD, conducted using the IVW method across 4 UK Biobank datasets (UKB-b-4657, UKB-b-1096, UKB-b-8875, and UKB-b-20124), revealed significant heterogeneity in the causal estimates derived from instrumental SNPs. Specifically, Cochran's Q statistic and the corresponding p-values indicated substantial heterogeneity across all datasets, with Q p-values ranging from 5.140×10^{-17} to 5.895×10^{-13} (Table 4). MR-Egger regression analysis further indicated the presence of directional pleiotropy. This conclusion was supported by statistically significant intercept p-values (p < 0.05) across all datasets, suggesting that genetic variants associated with *E2F2* may influence heel BMD through pathways other than the hypothesized causal mechanism (Table 5).

The MR-PRESSO analysis for horizontal pleiotropy, examining the causal effect of *E2F2* on heel BMD, consistently demonstrated a negative association across all datasets, as indicated by statistically significant p-values (p < 0.001) for the causal estimates. The global test p-value (p_{Glo}), which also remained below 0.001, further supported the overall significance of the observed association (Table 6).

Table 2. Overlap of DEGs between OI and *PLIN2*-related gene

Gene symbol	GeneID	p_{adj}	p-value	lfcSE	stat	log2FoldChange	baseMean	Description
TNFRSF19	55504	0.001	7.48E-10	0.446	–6.156	–2.745	1859.710	tumor necrosis factor (TNF) receptor superfamily member 19
E2F2	1870	0.001	0.001	0.43	–4.985	–2.143	62.150	E2F transcription factor 2
SH3PXD2A	9644	0.001	0.001	0.402	4.498	1.807	6808.540	SH3 and PX domains 2A
MMP11	4320	0.005	0.001	0.492	3.848	1.894	143.030	matrix metallopeptidase 11
TXNIP	10628	0.011	0.001	0.351	3.599	1.263	4511.310	thioredoxin interacting protein
PPP1R3G	648791	0.022	0.001	0.477	3.346	1.595	64	protein phosphatase 1 regulatory subunit 3G
ID3	3399	0.032	0.001	0.369	3.197	1.180	2679.070	inhibitor of DNA binding 3

Gene symbol – official gene symbol; GeneID – NCBI gene identifier; p_{adj} – adjusted p-value (Benjamini–Hochberg false discovery rate); p-value – raw p-value from Wald test; lfcSE – log2 fold change standard error; stat – Wald statistic; log2FoldChange – log2 fold change (experimental vs control); baseMean – mean of normalized counts across all samples; description – gene name and functional annotation; DEGs – differentially expressed genes; OI – osteogenesis imperfecta; NCBI – National Center for Biotechnology Information.

Table 3. Mendelian randomization (MR) analysis results for gene *E2F2* and bone mineral density (BMD) (method – inverse-variance weighted (IVW))

Datasets	nsnp	b	SE	p-value	Lower 95% CI	Upper 95% CI	OR	OR_LCI_95	OR_UCI_95
UKB-b-8875	29.000	–0.028	0.007	0.000	–0.041	–0.014	0.973	0.960	0.986
UKB-b-20124	29.000	–0.028	0.007	0.000	–0.042	–0.015	0.972	0.959	0.986
UKB-b-1096	29.000	–0.034	0.007	0.000	–0.049	–0.020	0.966	0.952	0.981
UKB-b-4657	29.000	–0.046	0.009	0.000	–0.063	–0.029	0.955	0.939	0.971

95% CI – 95% confidence interval; nsnp – number of single nucleotide polymorphisms; b – effect size; SE – standard error of effect size; OR – odds ratio; OR_LCI_95 – odds ratio lower boundary of 95% CI; OR_UCI_95 – odds ratio upper boundary of 95% CI.

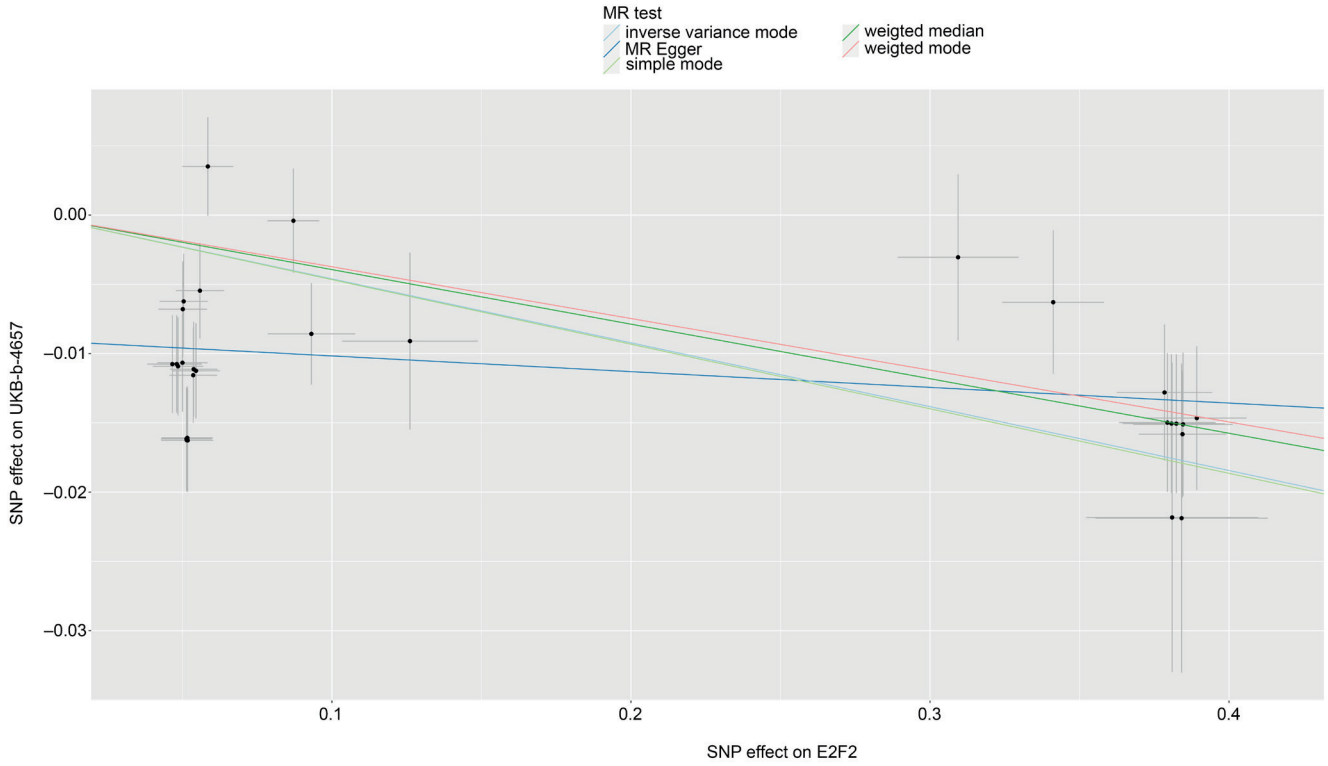

Fig. 2. Scatter plot of Mendelian randomization (MR) analysis for *E2F2* gene in the UKB-b-4657 dataset

Table 4. Heterogeneity analysis results for *E2F2* gene and heel BMD (method – inverse-variance weighted (IVW))

Datasets	Q	Q_df	Q_pval
UKB-b-4657	117.638	28	5.89582E-13
UKB-b-1096	86.642	28	6.57791E-08
UKB-b-8875	135.193	28	5.36E-16
UKB-b-20124	140.945	28	5.14E-17

Datasets – source genome-wide association study (GWAS) dataset identifier; BMD – bone mineral density; method – meta-analysis method (inverse-variance weighted); Q – Cochran's Q statistic; Q_df – degrees of freedom for Cochran's Q statistic; Q_pval – p-value for Cochran's Q statistic.

Table 5. MR-Egger analysis results for *E2F2* gene and heel BMD

Datasets	Egger_ intercept	SE	p-value
UKB-b-4657	−0.009	0.001	2.76E-07
UKB-b-1096	−0.008	0.001	3.52E-08
UKB-b-8875	−0.008	0.001	1.49E-09
UKB-b-20124	−0.008	0.001	1.36E-09

BMD – bone mineral density; SE – standard error of effect size.

Table 6. Horizontal pleiotropy analysis results for *E2F2* gene and heel BMD (MR analysis – raw)

Causal estimate	SD	t-stat	p-value	p_Glo
−0.046	0.009	−5.409	8.15E-06	<0.001
−0.035	0.007	−4.674	6.27E-05	<0.001
−0.028	0.007	−4.061	0.001	<0.001
−0.029	0.007	−4.066	0.001	<0.001

Causal estimate – causal effect estimate (beta coefficient) from MR analysis; SD – standard deviation; t-stat – t-statistic; p-value – p-value for the t-statistic; p_Glo – global p-value for horizontal pleiotropy; horizontal-pleiotropy – pleiotropic effects via independent pathways; MR – Mendelian randomization; raw – unadjusted analysis; BMD – bone mineral density.

Discussion

Our analysis revealed that *PLIN2* is a key gene influenced by MSC therapy in OI, exhibiting the most significant adjusted p-value among the identified DEGs. *PLIN2* plays a critical role in lipid droplet formation and storage, thereby influencing cellular lipid metabolism. Given the important role of lipids in cellular energy homeostasis and signaling, particularly in stem cell differentiation,[16] modulation of *PLIN2* by MSC therapy suggests a potential mechanism for altering the metabolic environment of bone-forming cells in OI. Furthermore, our findings indicate that *PLIN2* deficiency leads to substantial transcriptomic changes, underscoring its regulatory role in gene expression.[17]

The observed interplay between *PLIN2* and other DEGs, such as *TNFRSF19* and *E2F2*, which are significantly downregulated in patients with OI but upregulated under *PLIN2* knockdown conditions, highlights the complex relationship between lipid metabolism and bone homeostasis.[18] This finding suggests that *PLIN2* may act as a key regulator of the balance between osteogenic and adipogenic differentiation of MSCs.[19] Dysregulation of this balance is a hallmark of OI, in which impaired bone formation is often

accompanied by increased adipogenesis.[20] Accordingly, *PLIN2* emerges as a potential therapeutic target in MSC-based therapies for OI, as its modulation may help restore the balance between bone and adipose tissue formation, thereby improving bone matrix quality and osteoblast function. By targeting *PLIN2*, MSC therapy may not only enhance bone repair but also mitigate the adverse effects of excessive adipogenesis, potentially leading to more effective and sustained therapeutic outcomes in patients with OI.[21]

The complex interplay between bone and lipid metabolism is essential for maintaining skeletal health, particularly in conditions such as OI, in which metabolic imbalances may exacerbate disease pathology. Our findings underscore the important role of *PLIN2* in this context, consistent with the broader understanding of how lipid metabolism influences bone homeostasis.[22] The balance between osteogenesis and adipogenesis is tightly regulated, and its disruption can lead to impaired bone quality and increased fracture risk, as observed in OI.[23] The involvement of *PLIN2* in lipid metabolism suggests that it may play a key role in this process, potentially influencing the differentiation of MSCs into either osteogenic or adipogenic lineages. This is particularly relevant in OI, where a shift toward adipogenesis may impair effective bone formation.[24]

Research demonstrates that lipid metabolism is closely linked to bone homeostasis, and disruptions in this balance contribute to OI pathophysiology. For instance, the peroxisome proliferator-activated receptor (PPAR) signaling pathway, a key regulator of adipogenesis and osteogenesis, is influenced by lipid dynamics. Although *PLIN5* has been shown to modulate PPAR signaling in oxidative tissues,[25] the role of *PLIN2* in MSCs and adipocytes suggests that it may indirectly influence this pathway by regulating lipid availability. Specifically, *PLIN2* controls the availability of free fatty acids for energy metabolism compared to their storage within lipid droplets. In OI, where a shift toward adipogenesis is commonly observed, modulation of *PLIN2* may restrict lipid availability for adipogenic pathways and redirect metabolic resources toward osteogenesis. This balance is critical; by optimizing lipid partitioning, *PLIN2* may influence MSC lineage commitment, promoting differentiation toward osteoblasts rather than adipocytes. Bone remodeling involves a tightly regulated interplay between osteoblasts and osteoclasts, with lipid metabolism playing an important regulatory role in this process. Studies indicate that *PLIN5* deficiency alters the expression of osteoclastogenesis-related genes, such as *RANKL* and *OPG*, thereby affecting bone resorption. In contrast, the role of *PLIN2* in lipid storage within the bone microenvironment may indirectly stabilize osteoclast–osteoblast dynamics by modulating the availability of lipid-derived signaling molecules. These molecules, including fatty acids, are known to influence inflammation and oxidative stress – key contributors to the excessive bone fragility observed in OI.[26] Accordingly, the regulation of lipid reserves by *PLIN2* may support bone remodeling processes in the context of MSC-based therapy. The therapeutic relevance of *PLIN2* is further supported by its potential interaction with mitochondrial function, a critical component of energy metabolism in osteoblasts. Previous studies on *PLIN5* have highlighted its role in facilitating mitochondrial lipid utilization, thereby influencing adenosine triphosphate (ATP) production and ROS levels. In OI, mitochondrial dysfunction exacerbates oxidative stress and impairs osteogenesis, and *PLIN5* deficiency may further aggravate these effects. Although *PLIN2* is not directly associated with mitochondria, unlike *PLIN5*, it contributes to cytoplasmic lipid homeostasis and may indirectly support mitochondrial energy supply. This suggests that *PLIN2* could enhance MSC therapy by maintaining the lipid pool required for energy-intensive bone formation.

PLIN2 and *PLIN5*, both members of the perilipin family, exhibit distinct yet complementary roles in lipid and bone metabolism. *PLIN2* primarily regulates lipid droplet storage and hydrolysis in adipocytes and MSCs, whereas *PLIN5*, which is enriched in oxidative tissues, directs lipids toward mitochondria for β-oxidation. In the context of OI, the predominance of *PLIN2* in MSCs positions it as a key regulator of lipid availability, influencing whether these cells differentiate into osteoblasts or adipocytes. In contrast, the mitochondrial role of *PLIN5* supports energy metabolism in active osteoblasts and may help mitigate oxidative stress in OI. The interplay between these proteins is evident: *PLIN2* maintains the lipid reservoir, whereas *PLIN5* facilitates the utilization of these lipids for mitochondrial energy production.

One downstream consequence of the *PLIN2*-mediated inflammatory response is the upregulation of *TNFRSF19*, a member of the tumor necrosis factor (TNF) receptor superfamily. Evidence suggests that *TNFRSF19* expression is elevated under conditions of increased inflammation and cellular stress, potentially as a compensatory response.[27] In the context of *PLIN2* deficiency, increased availability of lipid-derived signaling molecules may stimulate *TNFRSF19*, which has been implicated in the regulation of cell proliferation and survival. Notably, GWAS have identified *TNFRSF19* among candidate genes associated with bone biology, alongside genes such as *PPARG* and *FBN2*.[28] The upregulation of *TNFRSF19* under *PLIN2* knockdown conditions suggests a potential role in osteoblast or pre-osteoblast proliferation, as *TNFRSF19* signaling may promote cell growth in response to inflammatory stimuli. In OI, where bone formation is already impaired, the interplay between *PLIN2* deficiency, inflammation, and *TNFRSF19* upregulation presents a complex scenario. On the one hand, enhanced osteoblast proliferation driven by *TNFRSF19* could potentially improve bone formation and counteract osteopenia. On the other hand, inflammation associated with increased lipid catabolism may offset these benefits by promoting osteoclast activity or impairing bone matrix quality.[29] This dual effect highlights the nuanced role of *PLIN2* as a regulator of both lipid metabolism and inflammatory tone in the bone microenvironment.

The observed upregulation of *TNFRSF19* under *PLIN2* knockdown conditions, as identified in our analysis, further supports its potential role in modulating bone cell dynamics, although the net outcome – whether anabolic or catabolic – likely depends on the balance between these competing processes. From a therapeutic perspective, these findings reinforce the importance of *PLIN2* as a target in MSC-based therapy for OI. Preservation of *PLIN2* function may help mitigate excessive lipolysis and inflammation, thereby limiting downstream activation of *TNFRSF19* and its uncertain effects on bone cell proliferation. Alternatively, if the proliferative effects of *TNFRSF19* prove beneficial, a combined strategy involving *PLIN2* modulation and targeted *TNFRSF19* agonists may enhance osteogenesis while maintaining control over inflammation. This hypothesis is supported by evidence that *TNFRSF19* contributes to bone biology and that *PLIN2* deficiency amplifies inflammatory signaling, underscoring the need for further studies to clarify these interactions.

The downregulation of *E2F2* in OI is consistent with its established role in bone and cartilage biology. *E2F2*, a transcription factor, is involved in cell cycle regulation and differentiation, with studies demonstrating its impact on skeletal tissues. For example, constitutive overexpression of *E2F2* inhibits chondrocyte differentiation and delays endochondral ossification.[30] These findings suggest that reduced *E2F2* expression in OI may reflect impaired chondrocyte or osteoblast differentiation, thereby contributing to the characteristic bone fragility observed in this disease. Conversely, the upregulation of *E2F2* under *PLIN2* knockdown conditions indicates that *PLIN2* may normally suppress *E2F2*, potentially maintaining a balance that favors osteogenesis over excessive proliferation or adipogenesis. This relationship is particularly relevant given the role of *PLIN2* in lipid metabolism, which influences MSC fate. Mechanistically, *E2F2* regulates osteogenesis and adipogenesis through its control of cell cycle progression and progenitor cell proliferation. By governing the transition from proliferation to differentiation, *E2F2* may define the temporal window during which MSCs commit to either osteogenic or adipogenic lineages. In OI, altered *E2F2* expression may disrupt this critical timing, thereby contributing to the imbalance in bone formation. Further evidence supports the therapeutic relevance of *E2F2* in skeletal disorders. In traumatic osteoarthritis, exosomes enriched with miR-125a-5p derived from MSCs target *E2F2* to suppress chondrocyte degeneration, suggesting a protective effect associated with *E2F2* downregulation.[31] Similarly, in osteoarthritis prevention, long noncoding RNA (lncRNA) PTS-1 inhibits *E2F2* via the miR-8085/*E2F2* axis, thereby promoting cartilage homeostasis.[32] These findings contrast with OI, in which *E2F2* downregulation may exacerbate bone defects, highlighting context-dependent effects. In fracture healing, quercetin modulates the miR-6089/*E2F2* axis to enhance osteoblast viability, proliferation, migration, and differentiation,[33] suggesting that increased *E2F2* expression may be beneficial in certain regenerative contexts. The upregulation of *E2F2* following *PLIN2* loss may therefore represent a compensatory mechanism in OI, aimed at enhancing osteoblast activity, although potentially at the expense of increased inflammation or dysregulated differentiation. Mendelian randomization analysis further elucidates the causal relationship between *E2F2* and bone health. Across 4 UK Biobank datasets, *E2F2* expression consistently demonstrated a negative association with heel BMD, with statistically significant p-values (from 1.116×10^{-7} to 6.073×10^{-5}) and effect estimates indicating that higher *E2F2* expression is associated with lower BMD. This finding is consistent with the inhibitory role of *E2F2* in chondrocyte differentiation and suggests that *E2F2* downregulation in OI may represent a compensatory mechanism aimed at preserving BMD, despite the osteopenic phenotype of the disease. However, the relatively small effect sizes (ORs: 0.955–0.973) indicate a modest magnitude of association. Moreover, the presence of significant heterogeneity and directional pleiotropy suggests that additional biological pathways may contribute to this relationship, thereby complicating its potential as a therapeutic target. This heterogeneity indicates that the causal effect estimates across instrumental variables are not fully consistent, which may arise from several factors. Potential explanations include residual confounding, differences in genetic architecture or environmental exposures within UK Biobank subpopulations, and horizontal pleiotropy, whereby SNPs influence the outcome through pathways independent of the exposure. Although MR-Egger and MR-PRESSO analyses were applied to detect and partially account for pleiotropy, residual pleiotropic effects or context-specific genetic interactions may still contribute to the observed variability. These findings suggest that the relationship between *E2F2* and BMD is multifactorial and likely governed by complex biological networks, rather than a single causal pathway. Accordingly, while the MR analysis supports a causal association, the magnitude and generalizability of this effect should be interpreted with caution. Future studies incorporating individual-level data or stratified population analyses are warranted to further elucidate the sources of heterogeneity and refine causal inference.

Mesenchymal stem cell therapy has attracted considerable attention as a potential treatment for OI due to its capacity to differentiate into osteoblasts, home to sites of injury, and exert paracrine effects on recipient tissues. These properties, together with a favorable safety profile, multilineage differentiation potential, low immunogenicity, and relative ease of manufacturing, position MSCs as a promising therapeutic platform for OI. Fetal-derived MSCs, in particular, exhibit enhanced osteogenic potential compared with adult-derived MSCs, as supported by preclinical and early clinical studies demonstrating improved bone formation and potential clinical benefits following intravenous administration.[34] Despite these advantages,

the therapeutic efficacy of MSC therapy in OI remains limited, with outcomes that are often modest, variable, and insufficiently durable, thereby hindering broader clinical translation. Evidence from multiple studies highlights these limitations. Although MSC therapy shows promise, including the superior bone-forming capacity of fetal MSCs, consistent and sustained clinical benefit has not been reliably achieved. More broadly, analyses of clinical-stage MSC therapies indicate that many fail to meet primary efficacy endpoints, suggesting that the intrinsic therapeutic potency of MSCs in humans may be lower than that predicted by preclinical models.[35] This discrepancy may be attributed to heterogeneity arising at multiple stages of cell therapy development, including cell sourcing, expansion, and delivery, all of which can undermine consistency and therapeutic efficacy. In osteoarthritis, MSC injections have been shown to improve clinical outcomes such as pain and joint function; however, magnetic resonance imaging (MRI) typically demonstrates only minimal structural improvement, suggesting that functional benefits may not translate into durable tissue repair.[36] Similarly, clinical trials of MSC therapy in osteoarthritis and rheumatoid arthritis report improvements in joint function, pain reduction, and quality of life without major adverse effects. Nevertheless, the magnitude of these benefits varies, and their long-term durability remains uncertain.[37] Additional challenges include age-related declines in MSC potency and the risk of immune responses associated with allogeneic MSCs.[38] Taken together, these findings indicate that although MSC therapy provides measurable benefits, its effects are often insufficient or unstable for addressing complex underlying pathologies such as OI. Accordingly, the inconsistent and modest efficacy of MSC-based approaches highlights the need to target key regulatory genes, including *PLIN2* and *E2F2*, to enhance osteogenesis and improve therapeutic stability, potentially through combinatorial strategies or MSC engineering.

Future directions and clinical relevance

Our study, primarily computational in nature, identifies *PLIN2* and *E2F2* as promising therapeutic targets in OI and suggests novel strategies to enhance MSC-based therapies by promoting osteogenesis and improving the stability of therapeutic outcomes. Future research should prioritize experimental validation of these findings using both in vitro and in vivo models to confirm the causal roles of *PLIN2* and *E2F2* in osteoblast and adipocyte differentiation. Further investigation into the molecular mechanisms by which modulation of these genes affects bone quality and fracture healing is warranted. Ultimately, these insights may facilitate the development of more targeted and effective MSC-based interventions for OI, supporting the advancement of personalized therapeutic approaches aimed at optimizing bone health and reducing disease burden.

Limitations of the study

Although our bioinformatics analysis provides valuable insights into the roles of *PLIN2* and *E2F2* in OI and their potential as therapeutic targets, it is inherently limited by its reliance on computational methods and publicly available datasets. The identification of DEGs and the causal associations inferred from MR analyses are based on statistical relationships rather than direct experimental evidence, and therefore should be interpreted with caution.

Factors such as sample heterogeneity, batch effects, and unmeasured confounders may have influenced the observed expression patterns and associations. Moreover, the upregulation of *E2F2* and *TNFRSF19* under *PLIN2* knockdown conditions, contrasted with their downregulation in OI, suggests a complex regulatory interplay that requires functional validation. Further experimental studies – including in vitro assays using MSC models, in vivo *PLIN2* knockout models, and histological analyses of bone formation – are necessary to elucidate the mechanistic role of *PLIN2* in lipid metabolism and the impact of *E2F2* on osteoblast differentiation and BMD. In addition, clinical translation will require well-designed trials to evaluate the feasibility and efficacy of targeting these pathways in MSC-based therapies, ensuring that bioinformatics-derived hypotheses are validated by robust experimental and clinical evidence.

Conclusions

Our study highlights the pivotal role of *PLIN2* in regulating lipid metabolism and MSC fate in OI, with *E2F2* emerging as a key downstream regulator that influences bone cell dynamics and BMD. The overlap of DEGs between OI and *PLIN2*-related profiles, together with MR evidence linking *E2F2* to skeletal traits, underscores their therapeutic relevance. Despite the promise of MSC-based therapies for OI, their efficacy remains modest and variable, characterized by transient clinical benefits, inconsistent outcomes, limited structural repair, and potential immunological challenges. These limitations highlight the need for more targeted and mechanistically informed approaches. We propose that *PLIN2* and *E2F2* represent critical molecular targets for enhancing MSC therapy. Precise modulation of *PLIN2* to optimize lipid metabolism, together with regulation of *E2F2*-mediated differentiation pathways, may improve osteogenic potential and promote more durable therapeutic effects. Although these findings provide a compelling framework, experimental and clinical validation is essential to translate bioinformatics-driven insights into effective therapeutic strategies, ultimately improving skeletal outcomes in patients with OI.

Data Availability Statement

All data used in this study are derived from publicly available repositories. This study analyzed datasets from the Gene Expression Omnibus (GEO; GSE157587, GSE214064, GSE186141) and UK Biobank genome-wide association study (GWAS) summary statistics (UKB-b-4657, UKB-b-1096, UKB-b-8875, UKB-b-20124).

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Xishun Wang https://orcid.org/0009-0001-9537-6953
Zhenjiang Liu https://orcid.org/0000-0003-2635-191X
Xinyong Hu https://orcid.org/0009-0009-0582-3878
Yinpeng Cui https://orcid.org/0000-0002-2070-6057

References

1. Görgün B, Yaşar NE, Bingöl İ, et al. Prevalence, number of fractures, and hospital characteristics among the pediatric population with osteogenesis imperfecta: Results from the nationwide registry of Türkiye. *J Pediatr Orthop B*. 2025;34(3):249–256. doi:10.1097/BPB.0000000000001192
2. Carbine KA, Larson MJ. Quantifying the presence of evidential value and selective reporting in food-related inhibitory control training: A *p*-curve analysis. *Health Psychol Rev*. 2019;13(3):318–343. doi:10.1080/17437199.2019.1622144
3. Andreotti JP, Lousado L, Magno LAV, Birbrair A. Hypothalamic neurons take center stage in the neural stem cell niche. *Cell Stem Cell*. 2017;21(3):293–294. doi:10.1016/j.stem.2017.08.005
4. Wagenlehner FME, Naber KG. Cefiderocol for treatment of complicated urinary tract infections. *Lancet Infect Dis*. 2019;19(1):22–23. doi:10.1016/S1473-3099(18)30722-9
5. Gargioli C, Slack JMW. Cell lineage tracing during *Xenopus* tail regeneration. *Development (Camb)*. 2004;131(11):2669–2679. doi:10.1242/dev.01155
6. Glatard T, Montagnat J, Pennec X. Medical image registration algorithms assesment: Bronze Standard application enactment on grids using the MOTEUR workflow engine. *Stud Health Technol Inform*. 2006;120:93–103. PMID:16823126.
7. Nandy A, Helderman RCM, Thapa S, et al. Enhanced fatty acid oxidation in osteoprogenitor cells provides protection from high-fat diet induced bone dysfunction. *J Bone Miner Res*. 2025;40(2):283–298. doi:10.1093/jbmr/zjae195
8. Cortini M, Ilieva E, Massari S, Bettini G, Avnet S, Baldini N. Uncovering the protective role of lipid droplet accumulation against acid-induced oxidative stress and cell death in osteosarcoma. *Biochim Biophys Acta Mol Basis Dis*. 2025;1871(2):167576. doi:10.1016/j.bbadis.2024.167576
9. Angelini G, Panunzi S, Pompili M, et al. Performance of noninvasive tests for metabolic dysfunction-associated steatohepatitis and liver fibrosis resolution after bariatric surgery. *Clin Chem*. 2025;71(3):406–417. doi:10.1093/clinchem/hvae208
10. Deng X, Liu H, Zhao W, et al. Expression of AMPK and *PLIN2* in the regulation of lipid metabolism and oxidative stress in bitches with open cervix pyometra. *BMC Vet Res*. 2025;21(1):164. doi:10.1186/s12917-025-04622-1
11. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550. doi:10.1186/s13059-014-0550-8
12. Magno R, Maia AT. gwasrapidd: An R package to query, download and wrangle GWAS catalog data. *Bioinformatics*. 2020;36(2):649–650. doi:10.1093/bioinformatics/btz605
13. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife*. 2018;7:e34408. doi:10.7554/eLife.34408
14. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018;50(5):693–698. doi:10.1038/s41588-018-0099-7
15. Zhu Z, Zhang F, Hu H, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. *Nat Genet*. 2016;48(5):481–487. doi:10.1038/ng.3538
16. Tsai TH, Chen E, Li L, et al. The constitutive lipid droplet protein *PLIN2* regulates autophagy in liver. *Autophagy*. 2017;13(7):1130–1144. doi:10.1080/15548627.2017.1319544
17. Wu Y, Chen K, Li L, et al. *PLIN2*-mediated lipid droplet mobilization accelerates exit from pluripotency by lipidomic remodeling and histone acetylation. *Cell Death Differ*. 2022;29(11):2316–2331. doi:10.1038/s41418-022-01018-8
18. Kaushik S, Cuervo AM. AMPK-dependent phosphorylation of lipid droplet protein *PLIN2* triggers its degradation by CMA. *Autophagy*. 2016;12(2):432–438. doi:10.1080/15548627.2015.1124226
19. Luo H, She X, Zhang Y, et al. *PLIN2* promotes lipid accumulation in ascites-associated macrophages and ovarian cancer progression by HIF1α/SPP1 signaling. *Adv Sci (Weinh)*. 2025;12(12):2411314. doi:10.1002/advs.202411314
20. Li Y, Khanal P, Norheim F, et al. *PLIN2* deletion increases cholesteryl ester lipid droplet content and disturbs cholesterol balance in adrenal cortex. *J Lipid Res*. 2021;62:100048. doi:10.1016/j.jlr.2021.100048
21. Mardani I, Tomas Dalen K, Drevinge C, et al. *PLIN2*-deficiency reduces lipophagy and results in increased lipid accumulation in the heart. *Sci Rep*. 2019;9(1):6909. doi:10.1038/s41598-019-43335-y
22. Kong L, Zhao H, Wang F, et al. Endocrine modulation of brain-skeleton axis driven by neural stem cell-derived perilipin 5 in the lipid metabolism homeostasis for bone regeneration. *Mol Ther*. 2023;31(5):1293–1312. doi:10.1016/j.ymthe.2023.02.004
23. Zhang J, Hu W, Zou Z, et al. The role of lipid metabolism in osteoporosis: Clinical implication and cellular mechanism. *Genes Dis*. 2024;11(4):101122. doi:10.1016/j.gendis.2023.101122
24. Peng H, Hu B, Xie LQ, et al. A mechanosensitive lipolytic factor in the bone marrow promotes osteogenesis and lymphopoiesis. *Cell Metab*. 2022;34(8):1168–1182.e6. doi:10.1016/j.cmet.2022.05.009
25. Xiao H, Li W, Qin Y, et al. Crosstalk between lipid metabolism and bone homeostasis: Exploring intricate signaling relationships. *Research (Wash D C)*. 2024;7:0447. doi:10.34133/research.0447
26. Dai B, Xu J, Li X, et al. Macrophages in epididymal adipose tissue secrete osteopontin to regulate bone homeostasis. *Nat Commun*. 2022;13(1):427. doi:10.1038/s41467-021-27683-w
27. Qiu W, Hu Y, Andersen TE, et al. Tumor necrosis factor receptor superfamily member 19 (*TNFRSF19*) regulates differentiation fate of human mesenchymal (stromal) stem cells through canonical Wnt signaling and C/EBP. *J Biol Chem*. 2010;285(19):14438–14449. doi:10.1074/jbc.M109.052001
28. Pei YF, Liu L, Liu TL, et al. Joint association analysis identified 18 new loci for bone mineral density. *J Bone Miner Res*. 2019;34(6):1086–1094. doi:10.1002/jbmr.3681
29. Mekchay S, Pothakam N, Norseeda W, et al. Association of *IFNA16* and *TNFRSF19* polymorphisms with intramuscular fat content and fatty acid composition in pigs. *Biology (Basel)*. 2022;11(1):109. doi:10.3390/biology11010109
30. Scheijen B, Bronk M, Van Der Meer T, Bernards R. Constitutive E2F1 overexpression delays endochondral bone formation by inhibiting chondrocyte differentiation. *Mol Cell Biol*. 2003;23(10):3656–3668. doi:10.1128/MCB.23.10.3656-3668.2003
31. Xia Q, Wang Q, Lin F, Wang J. miR-125a-5p-abundant exosomes derived from mesenchymal stem cells suppress chondrocyte degeneration via targeting *E2F2* in traumatic osteoarthritis. *Bioengineered*. 2021;12(2):11225–11238. doi:10.1080/21655979.2021.1995580
32. Ma C, Chen Q, Wei YF, et al. LncRNA PTS-1 protects against osteoarthritis through the *miR-8085/E2F2* axis. *J Inflamm Res*. 2025;18:347–366. doi:10.2147/JIR.S496185

33. Dong R, Liu MY, Zhu GB, Tan KM, Wang YQ, Li L. Modulation of the microRNA-6089/E2F transcription factor2 axis by querceting: Implications for osteoblast viability, proliferation, migration, and osteogenic differentiation in fracture healing. *J Physiol Pharmacol*. 2024;75(2):173–183. doi:10.26402/jpp.2024.2.06
34. Götherström C, Walther-Jallow L. Stem cell therapy as a treatment for osteogenesis imperfecta. *Curr Osteoporos Rep*. 2020;18(4):337–343. doi:10.1007/s11914-020-00594-3
35. Levy O, Kuai R, Siren EMJ, et al. Shattering barriers toward clinically meaningful MSC therapies. *Sci Adv*. 2020;6(30):eaba6884. doi:10.1126/sciadv.aba6884
36. Fuggle NR, Cooper C, Oreffo ROC, et al. Alternative and complementary therapies in osteoarthritis and cartilage repair. *Aging Clin Exp Res*. 2020;32(4):547–560. doi:10.1007/s40520-020-01515-1
37. Hwang JJ, Rim YA, Nam Y, Ju JH. Recent developments in clinical applications of mesenchymal stem cells in the treatment of rheumatoid arthritis and osteoarthritis. *Front Immunol*. 2021;12:631291. doi:10.3389/fimmu.2021.631291
38. Bagge J, Freude K, Lindegaard C, Holst B, Hölmich P. Pros and cons of using autologous versus allogenic stem cells for the treatment of osteoarthritis [in Danish]. *Ugeskr Læger*. 2024;186(1):V06230423. doi:10.61409/V06230423

Prognostic model of endoplasmic reticulum stress-related lncRNAs in lung adenocarcinoma: Construction and validation using WGCNA

Haiyang Li[1,A], Zhenshan Zhao[2,D,F], Jing Li[1,B], Yao Rong[1,C], Aimin Zheng[3,B,C], Menghui Hao[2,C]

[1] Department of Medical Oncology, Kailuan General Hospital, Tangshan, China
[2] Department of Thoracic Surgery, Kailuan General Hospital, Tangshan, China
[3] Oncology Department of Integrated Chinese and Western Medicine, Tangshan People's 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(5):847–861

Address for correspondence
Zhenshan Zhao
E-mail: zhaozhenshan1978@126.com

Funding sources
None declared

Conflict of interest
None declared

Received on March 8, 2025
Reviewed on May 5, 2025
Accepted on August 18, 2025

Published online on May 26, 2026

Abstract

Background. Lung adenocarcinoma (LUAD) ranks among the deadliest malignancies worldwide. The endoplasmic reticulum (ER) stress response plays a critical role in the pathogenesis of various cancers, and long non-coding RNAs (lncRNAs) are known for their regulatory roles in gene expression and disease progression.

Objectives. To construct and validate a prognostic model based on ER stress-related lncRNAs in LUAD.

Materials and methods. The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) databases were used. Utilizing the Molecular Signatures Database (MSigDB), we identified ER stress-related mRNAs and lncRNAs. Weighted gene co-expression network analysis (WGCNA) was employed to identify genes associated with LUAD prognosis. An lncRNA-based prognostic risk scoring model was constructed using univariate and least absolute shrinkage and selection operator (LASSO) regression analyses and independently validated. Consensus clustering analysis was applied to define risk subgroups, optimizing the risk scoring system. The model's performance was evaluated using receiver operating characteristic (ROC) curves and nomograms. Differentially expressed gene (DEG) and enrichment analyses were performed to investigate the biological relevance of the risk score. Additionally, the relationships between risk scores, immune infiltration, and the tumor microenvironment (TME) were explored.

Results. Using WGCNA, we successfully identified genes strongly associated with ER stress in LUAD prognosis. A prognostic model comprising 13 signature genes was developed, demonstrating robust discrimination between high- and low-risk patients, with the high-risk group exhibiting reduced overall survival (OS). The model's predictive accuracy was confirmed through Kaplan–Meier and ROC analyses. Correlation analysis between risk scores and immune infiltration indicated that the model reflects the immune landscape of LUAD. Subgroup analysis using consensus clustering (C1 and C2) revealed more pronounced differences in immune cell dynamics than the binary risk score classification alone.

Conclusions. This study introduces a novel prognostic model based on the co-expression of ER stress-related lncRNAs in LUAD.

Key words: lung adenocarcinoma, immune infiltration, prognostic model, endoplasmic reticulum stress, long non-coding RNAs

Cite as
Li H, Zhao Z, Li J, Rong Y, Zheng A, Hao M. Prognostic model of endoplasmic reticulum stress-related lncRNAs in lung adenocarcinoma: Construction and validation using WGCNA. *Adv Clin Exp Med*. 2026;35(5):847–861. doi:10.17219/acem/209665

DOI
10.17219/acem/209665

Highlights

- A novel prognostic model based on endoplasmic reticulum (ER) stress-related long non-coding RNAs (lncRNAs) was developed for lung adenocarcinoma (LUAD).
- A 13-lncRNA risk signature effectively stratifies LUAD patients into high- and low-risk groups with distinct survival outcomes.
- DNAJC5B is associated with immune cell infiltration and ER stress regulation, highlighting potential therapeutic relevance in LUAD.
- The prognostic model demonstrated robust predictive performance (AUC > 0.7) across training, testing, and validation cohorts.
- ER stress-related lncRNA signatures may support prognostic assessment and personalized immunotherapy in LUAD.

Background

Lung adenocarcinoma (LUAD), the most common histological subtype of non-small cell lung cancer (NSCLC), is one of the most prevalent malignancies worldwide, accounting for approx. 40% of all lung cancer cases.[1–3] According to the World Health Organization (WHO), lung cancer remains the leading cause of cancer-related mortality globally, with adenocarcinoma representing a substantial proportion of these cases.[3] Lung adenocarcinoma is characterized by marked molecular heterogeneity and biological complexity, which contribute to its poor prognosis, particularly in patients diagnosed at advanced stages. The 5-year survival rate for LUAD remains low, often below 20%, highlighting the urgent need for improved diagnostic and therapeutic strategies. Therefore, a deeper understanding of the molecular mechanisms underlying LUAD and the identification of reliable biomarkers for early detection, targeted intervention, and personalized management are of critical importance.[4]

The endoplasmic reticulum (ER) is a critical intracellular organelle responsible for maintaining protein homeostasis through the regulation of protein synthesis, folding, and transport.[5,6] Endoplasmic reticulum stress occurs when ER protein-folding capacity is disrupted, resulting in the accumulation of misfolded proteins. This stress response activates a series of signaling pathways, including the unfolded protein response (UPR), which aims to restore cellular homeostasis.[7] In the context of LUAD, ER stress has been shown to be closely associated with malignant cellular behaviors, including invasion, metastasis, and resistance to anticancer therapies.[8–10] For example, studies have demonstrated that ER stress can promote epithelial–mesenchymal transition (EMT) in LUAD cells, thereby enhancing their invasive and metastatic potential. Additionally, ER stress has been implicated in the development of therapeutic resistance, representing a major challenge in cancer treatment. Therefore, targeting ER stress has emerged as a promising therapeutic strategy for LUAD.[11]

Long non-coding RNAs (lncRNAs) have attracted increasing attention in recent years due to their critical roles in diverse biological processes, including cellular differentiation, development, and gene expression regulation.[12] In LUAD, dysregulated lncRNA expression has been closely associated with tumor progression, metastasis, and patient prognosis.[13–16] For example, specific lncRNAs have been identified as potential oncogenes or tumor suppressors in LUAD, influencing cell proliferation, apoptosis, and migration. Moreover, lncRNAs can interact with ER stress-related signaling pathways, thereby modulating the biological behavior of tumor cells.[17] This crosstalk between lncRNAs and ER stress pathways suggests that ER stress-related lncRNAs may serve not only as promising biomarkers for LUAD but also as potential therapeutic targets.

Recent advances in molecular biology have revealed that the interplay between ER stress and lncRNAs is more complex than previously recognized. For example, the lncRNA *HOTAIR* has been reported to interact with ER stress-related pathways, promoting tumor progression in LUAD through modulation of key genes involved in the UPR.[18] Similarly, the lncRNA *MALAT1* has been implicated in enhancing the metastatic potential of LUAD cells by stabilizing ER stress-related transcripts.[19] These findings further underscore the potential of lncRNAs as both prognostic biomarkers and therapeutic targets in LUAD.

Furthermore, the tumor microenvironment (TME) plays a critical role in the progression of LUAD and in determining treatment response. Endoplasmic reticulum stress has been shown to modulate the TME by influencing the recruitment, activation, and function of immune cells, including macrophages and T lymphocytes.[20] This interaction between ER stress and the TME further complicates the therapeutic landscape of LUAD and highlights the need for a comprehensive understanding of the underlying molecular mechanisms. The complex interplay among ER stress, lncRNAs, and the TME in LUAD underscores the importance of further investigation into their roles in tumor biology. A deeper understanding of these interactions may reveal novel therapeutic targets and contribute to improved clinical outcomes in LUAD.

Objectives

Although previous studies have explored the roles of ER stress and ER stress-related lncRNAs in LUAD, their interactions and specific prognostic applications in LUAD remain unclear. This study aimed to construct an ER stress-related prognostic risk model through comprehensive analysis of gene expression data from LUAD patients, with particular emphasis on lncRNA expression patterns. We anticipate that this study will provide novel insights and potential biomarkers for personalized prognostic assessment and therapeutic decision-making in LUAD.

Material and methods

Data processing

This study utilized data from The Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov) and the Genotype-Tissue Expression (GTEx) project (https://gtexportal.org). In addition, single-cell validation analyses were performed using TISCH2 (http://tisch.comp-genomics.org), which includes 17 single-cell sequencing datasets. TISCH2 is a single-cell transcriptomic database focused on the TME, providing detailed cell-type annotations. As LUAD was not available as a specific category in the TISCH2 database, NSCLC was selected as the closest alternative for the initial validation analysis. This approach enabled validation using the available single-cell transcriptomic data relevant to the lung cancer microenvironment.

The TCGA cohort of patients with LUAD included transcriptomic data and detailed clinical information for 600 patients. All patients underwent surgical resection and had relatively complete follow-up data, with a median follow-up duration of 6 years. The patients ranged in age from 30 to 90 years, with a male-to-female ratio of approx. 1.8:1. The pathological stage distribution included 279 patients with stage I disease, 124 with stage II, 100 with stage IIA, 85 with stage III, and 26 with stage IV disease.

Identification of ER stress-related lncRNAs

A keyword search for "endoplasmic reticulum stress" identified 419 ER stress-related genes. Pearson's correlation analysis was subsequently performed between these ER stress-related genes and 16,876 lncRNAs in the expression matrix, identifying 3,611 lncRNAs significantly associated with ER stress.

Weighted correlation network analysis

Weighted correlation network analysis (WGCNA; v. 1.73) was performed to identify gene modules associated with LUAD prognosis. Selection of an appropriate soft-thresholding power is a critical step in WGCNA, as it directly influences the scale-free topology of the constructed network. To determine the optimal soft threshold, a systematic evaluation of candidate soft-thresholding powers ranging from 1 to 20 was performed. For each candidate value, the scale-free topology fit index (R^2) was calculated to assess the degree of conformity between the network connectivity distribution and the ideal scale-free topology. A soft-thresholding power of 16 was selected, at which the scale-free topology fit index reached 0.80, indicating a satisfactory approximation to a scale-free network model. This threshold was chosen based on the balance between module stability, intramodular connectivity, and reduction of background noise, thereby improving the identification of prognostically relevant gene modules in LUAD.

Construction and validation of the prognostic ER stress-related lncRNA risk score

For construction of the prognostic risk model, differential expression analysis between LUAD and control samples was first performed using the limma package (v. 3.50.3; https://bioconductor.org/packages/limma) in R to identify significantly differentially expressed genes (DEGs). Differentially expressed genes were defined using the thresholds of |log2 fold change| >1 and Benjamini–Hochberg adjusted $p < 0.05$. Subsequently, WGCNA was applied to the gene expression dataset to identify gene modules significantly associated with LUAD prognosis. WGCNA enables the identification of potential disease-related biomarkers by constructing gene co-expression networks and clustering genes with similar expression patterns into biologically relevant modules.

After identifying disease-related gene modules, intersecting analysis was performed between these modules and the previously identified DEGs. This step was used to further refine the candidate gene set by focusing on genes that were both differentially expressed and associated with disease-relevant molecular modules. Subsequently, univariate Cox regression analysis was performed to identify genes significantly associated with patient prognosis among the intersecting genes. Univariate Cox regression is a standard survival analysis method used to evaluate the association between gene expression levels and patient survival outcomes.

To further optimize the prognostic model and reduce the risk of overfitting, least absolute shrinkage and selection operator (LASSO) Cox regression analysis was performed to select prognostically relevant genes and estimate their coefficients. LASSO regression is a regularization-based method suitable for variable selection in high-dimensional datasets. The analysis was conducted using the glmnet package in R (https://CRAN.R-project.org/package=glmnet), and the optimal penalty parameter (λ) was determined through 10-fold cross-validation

by minimizing the partial likelihood deviance. This approach improved model robustness and generalizability.

Validation of prognostic model and consensus clustering

Based on gene expression profiles following risk stratification, consensus clustering analysis was performed on the TCGA-LUAD dataset using the ConsensusClusterPlus package (https://www.bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html). Principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) were used to visualize the distribution of clustered samples. A heatmap was generated to illustrate the distribution of LUAD patients across different clusters, together with survival information and stratification into high- and low-risk groups. A nomogram was constructed to provide a visual representation of the clinical applicability of the prognostic model. In addition, subgroup analyses based on clinical characteristics were performed to evaluate the prognostic value of the risk score and consensus clustering within specific patient subgroups.

Enrichment and TME analysis

Immune cell infiltration in LUAD samples was initially evaluated using the CIBERSORT algorithm (R script v. 1.03) (https://cibersort.stanford.edu) with 1,000 permutations and a significance threshold of $p < 0.05$, in combination with the IOBR (Immuno-Oncology Biological Research) package (v. 0.99.9; https://CRAN.R-project.org/package=IOBR) in R to estimate the relative abundance of immune cell populations based on RNA transcriptomic data. This approach enabled quantitative assessment of immune cell infiltration across individual samples. For TME analysis, the ESTIMATE algorithm implemented in the IOBR package was used to assess stromal and immune components within the tumor tissue. Default gene signatures were applied for immune and stromal cell estimation, and immune, stromal, and ESTIMATE scores were calculated for each TCGA-LUAD sample. Patients were stratified into high- and low-risk groups, as well as consensus clustering subgroups, using the median risk score as the cutoff. Differences in stromal, immune, and ESTIMATE scores between groups were assessed using the Wilcoxon rank-sum test.

Gene mutation analysis

Somatic mutation data for TCGA-LUAD patients were obtained from the TCGA database. A waterfall plot was generated to visualize the mutation landscape of frequently mutated genes in the high- and low-risk groups. In addition, tumor mutation burden (TMB) was calculated for each TCGA-LUAD patient using the maftools R package (v. 3.21; https://bioconductor.org/packages/release/bioc/html/maftools.html). Tumor mutation burden was defined as the number of somatic mutations per megabase (Mb) of sequenced genomic region and was subsequently used for comparative and correlation analyses.

LncRNA-related target genes

Potential target genes of the model lncRNAs were identified using RNAct (http://rnact.crg.eu) and systematically compiled for downstream analysis. An interaction network was constructed and hub gene analysis was performed using Cytoscape (https://cytoscape.org). Core target genes associated with the model lncRNAs were subsequently subjected to differential expression analysis across risk score groups and consensus clustering subgroups.

Statistical analyses

Statistical analyses and data visualization were performed using R v. 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). Survival analyses were conducted using the Kaplan–Meier method, and differences between groups were compared using the log-rank test. Associations between risk scores and clinical characteristics were evaluated using Cox proportional hazards regression analysis with the survival package (v. 3.2-11; https://github.com/therneau/survival) in R to assess the independent prognostic value of the risk score. During model construction and validation, the proportional hazards (PH) assumption was assessed for all Cox regression models, including univariable and LASSO Cox regression analyses. The PH assumption was evaluated using the Schoenfeld residuals test, and the results confirmed that the included variables satisfied this assumption (Supplementary Fig. 1).[21] Statistical tests were selected according to data distribution and variance characteristics. For comparisons between 2 groups, Student's t-test was applied when data met assumptions of normality and homogeneity of variance; Welch's t-test was used when normality was satisfied but variances were unequal; otherwise, the Wilcoxon rank-sum test was applied. A $p < 0.05$ was considered statistically significant.

Results

Model construction

Differential expression analysis and WGCNA were performed on LUAD and normal control samples from the TCGA and GTEx datasets (Fig. 1A,B). The intersection between differentially expressed genes and genes from the blue and gray WGCNA modules was selected for univariate Cox regression analysis (Fig. 1D), yielding candidate genes for prognostic model construction. Subsequently, LASSO Cox regression analysis was performed to construct the ER stress-related lncRNA prognostic model (Fig. 1E,F), with the following risk score formula:

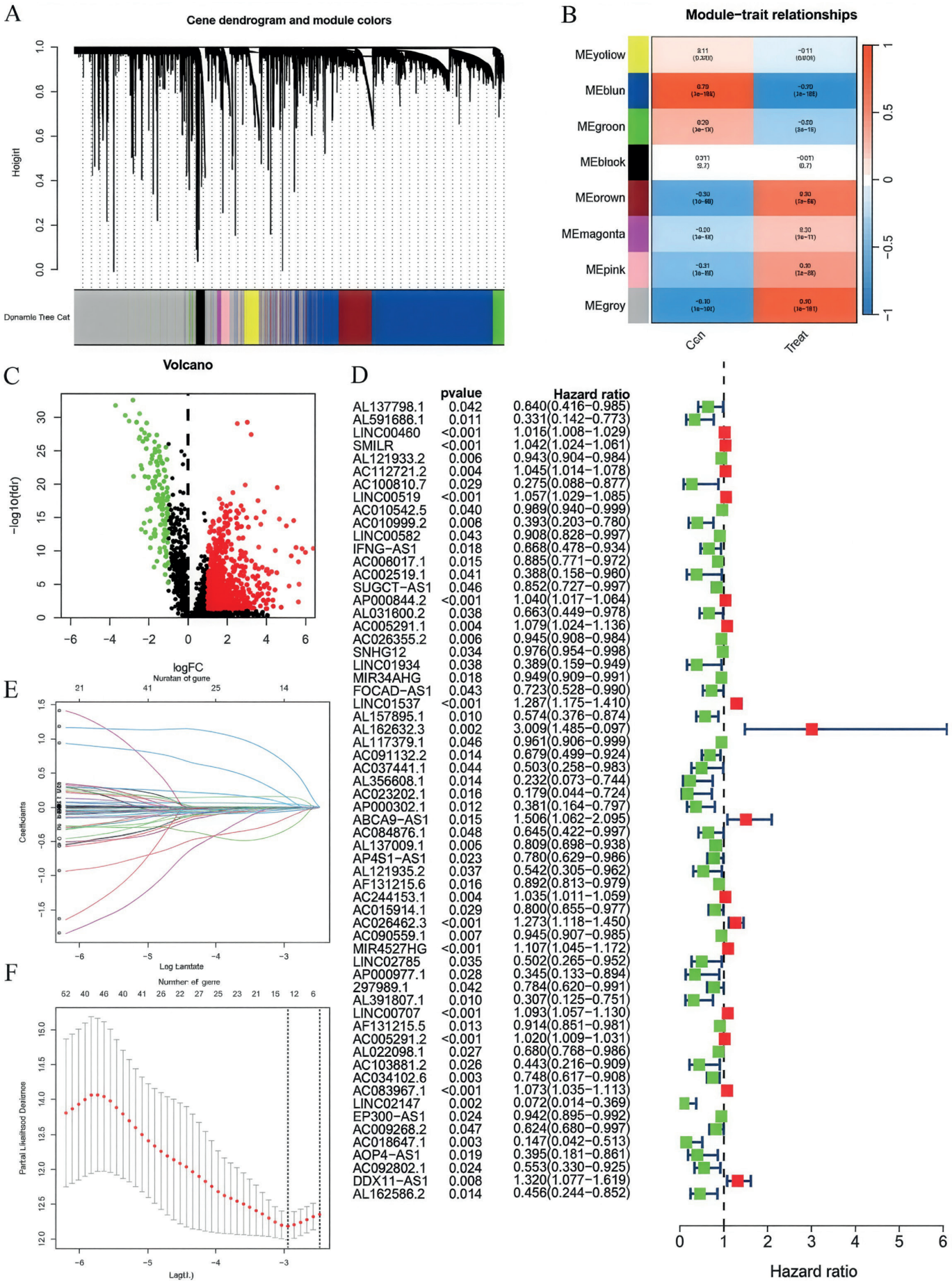

Fig. 1. Endoplasmic reticulum (ER) stress-related long non-coding RNA (lncRNA) model prognostic model. A. Tree-like topology; B. Module-trait correlation heatmap. The horizontal axis represents clinical characteristics, and the vertical axis represents the names of WGCNA modules; C. Volcano plot. Red represents high expression in tumors, and green represents low expression in tumors; D. Univariate Cox forest plot, red represents hazard ratio (HR) >1 and green represents HR < 1; E. Log change coefficients plotted with (least absolute shrinkage and selection operator (LASSO) Cox regression in 10-fold cross-validation; F. Partial likelihood deviation plotted with LASSO Cox regression in 10-fold cross-validation

Risk score =
(*LINC00460* × 0.0006) + (*AC010999.2* × −0.1509) + (*AC026355.2* × −0.0104) + (*LINC01537* × 0.1545) + (*AL162632.3* × 0.6158) + (*ABCA9-AS1* × 0.1352) + (*AL137009.1* × −0.0158) + (*AC026462.3* × 0.0417) + (*AC090559.1* × −0.0023) + (*LINC00707* × 0.0481) + (*AC034102.8* × −0.0575) + (*AC018647.1* × −0.3211) + (*AC092802.1* × −0.0283).
All lncRNAs included in the final prognostic model satisfied the proportional hazards assumption based on Schoenfeld residual testing.

Consensus clustering

Based on the model genes and risk scores, we conducted consensus clustering analysis on LUAD samples. First, the patients were divided into different clusters (K = 2–9), and the optimal number of clusters was determined to be K = 2 based on the consensus matrix (Fig. 2A). The consensus cumulative distribution function (CDF) curve was also plotted (Fig. 2B). The analysis showed that patients with LUAD could be divided into 2 risk groups, C1 and C2. The infiltration level of immunosuppressive cells (such as regulatory T cells, myeloid-derived suppressor cells, and tumor-associated macrophages) in the TME of group C1 (high-risk group) was relatively high. These cells can inhibit the antitumor immune response and promote tumor immune escape. Therefore, group C1 may be associated with a poorer survival prognosis. Conversely, the infiltration level of immunosuppressive cells in group C2 (low-risk group) was lower, suggesting that there may be a more active antitumor immune response in the TME. Effector T cells can recognize and attack tumor cells more effectively, thereby inhibiting tumor growth and metastasis. To visually demonstrate the relationship between the risk score and the clustering results, we plotted a Sankey plot (Fig. 2C) and verified using the Kaplan–Meier curve that the survival time of patients in group C2 was longer (Fig. 2D). Furthermore, the stability and repeatability of the clustering results were confirmed through PCA and t-SNE analysis (Fig. 2E,F).

Model validation and comparison

The samples were randomly divided into a test group and a training group. There were survival differences according to risk scores across the entire cohort, the test set, and the training set, with lower scores indicating better survival outcomes (Fig. 3A–C). Time-dependent receiver operating characteristic (ROC) analysis demonstrated robust predictive accuracy. The area under the ROC curve (AUC) for the risk scores with clinical relevance, as well as the AUCs at 1, 3, and 4 years, were consistent (Fig. 3D,E).

Functional enrichment analysis

Functional enrichment analysis was performed on differential genes between the risk score and consensus clustering groups. Gene Ontology (GO) analysis related to the risk score mainly involved cell movement and ciliary movement. Among these, "cell motility" and "cilia motility" were significantly enriched biological processes that play a key role in LUAD invasion and metastasis (Fig. 4A), while GO analysis related to consensus clustering mainly involved muscle contraction and fiber contraction (Fig. 4B). Kyoto Encyclopedia of Genes and Genomes (KEGG; https://www.genome.jp/kegg) enrichment analysis and gene set enrichment analysis (GSEA) related to the risk score were enriched only in cell movement; therefore, these results are not shown in the figures. The KEGG analysis for consensus clustering mainly involved vascular contraction (Fig. 4C). GSEA for consensus clustering mainly involved the formation and regulation of the stromal group (Fig. 4D).

Immune-related analysis

The risk score and consensus clustering showed similar profiles in immune cell infiltration and immune cell function, primarily associated with T cells and macrophages (Fig. 5A,B,D,E). By integrating checkpoint analysis with risk scores and consensus clustering, together with the analysis of immune cells and their functions, the consensus clustering groups appeared to be more immunologically sensitive (Fig. 5C,F).

Mutation analysis

An in-depth examination of the mutational profile in relation to the risk score was performed. The mutational spectrum revealed the top 20 most frequently mutated genes, highlighting a pronounced disparity in mutation rates between the high- and low-risk cohorts (Fig. 6A). A statistically significant difference in mutation rates was observed between these groups ($p = 0.039$), with the high-risk group exhibiting a notably elevated TMB score compared to the low-risk group (Fig. 6B). Moreover, a significant positive correlation was identified between TMB and risk scores (correlation coefficient $R = 0.087$, $p = 0.033$), suggesting a direct association between the 2 metrics (Fig. 6C).

Clinical and functional exploration of model genes

Pairwise sample analysis was conducted on the model genes, revealing expression differences for all core genes in the model except *AC010999.2*, *AC026462.3*, and *AC034102.8* (Fig. 7A). Prognostic analysis of the model genes over 1–5 years indicated a close relationship between

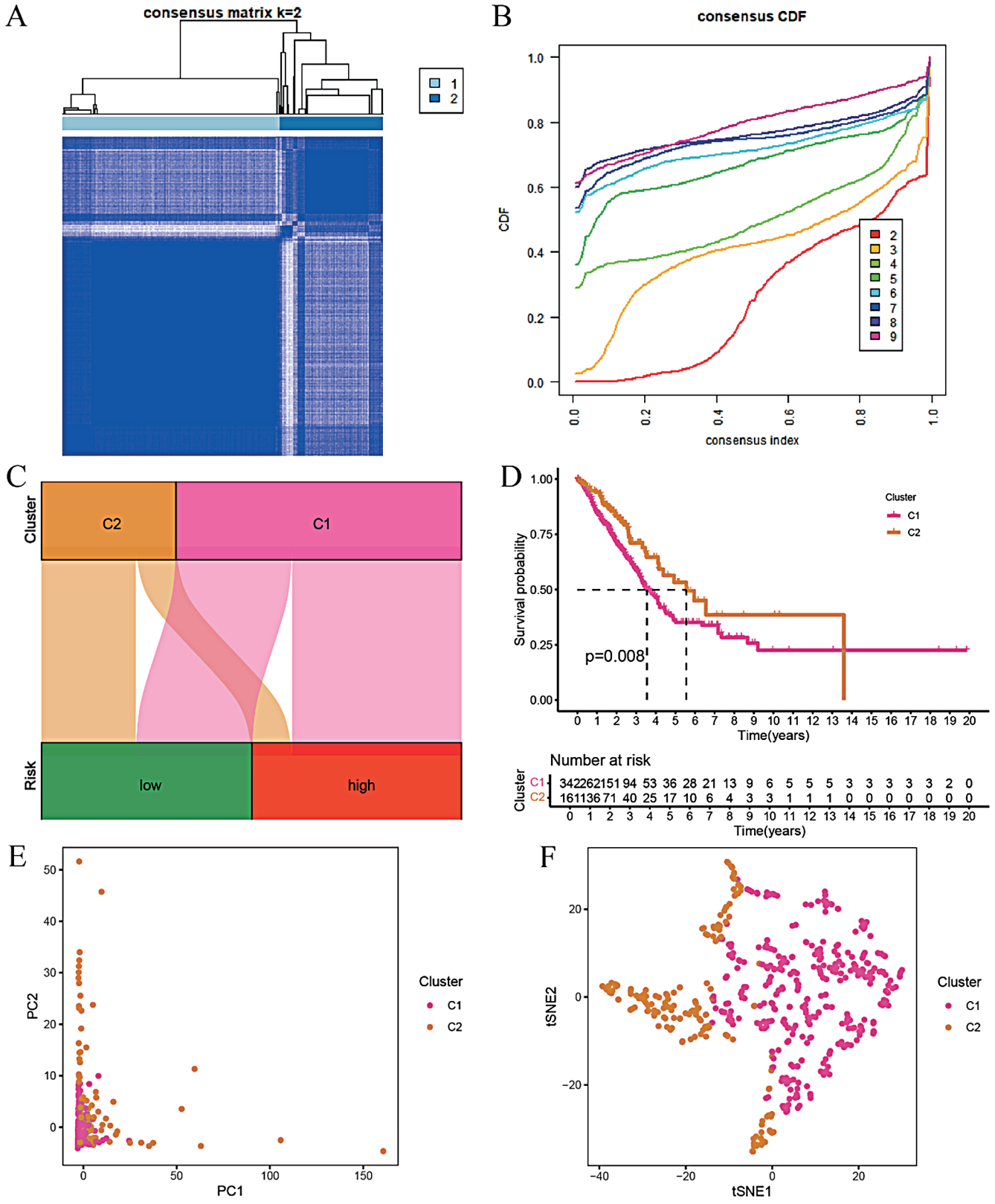

Fig. 2. Identifying risk clusters through consensus clustering. A. Consensus clustering matrix at K = 2; B. Cumulative distribution function (CDF) curves for clustering from k = 2 to 9; C. Sankey diagram; D. Survival analysis of lung adenocarcinoma (LUAD) samples divided into 2 clusters; E. Principal component analysis (PCA) plot for the 2 clusters; F. T-distributed stochastic neighbor embedding (tSNE) plot for the 2 clusters

the model genes and prognosis (Fig. 7B). Immunoinfiltration analysis showed that the model genes were involved in immune cell production and immune cell function (Fig. 7C,D). Based on TCGA data, we further analyzed the expression of the hub target genes in the risk scoring and consensus clustering groups (Fig. 7E,F). Combining

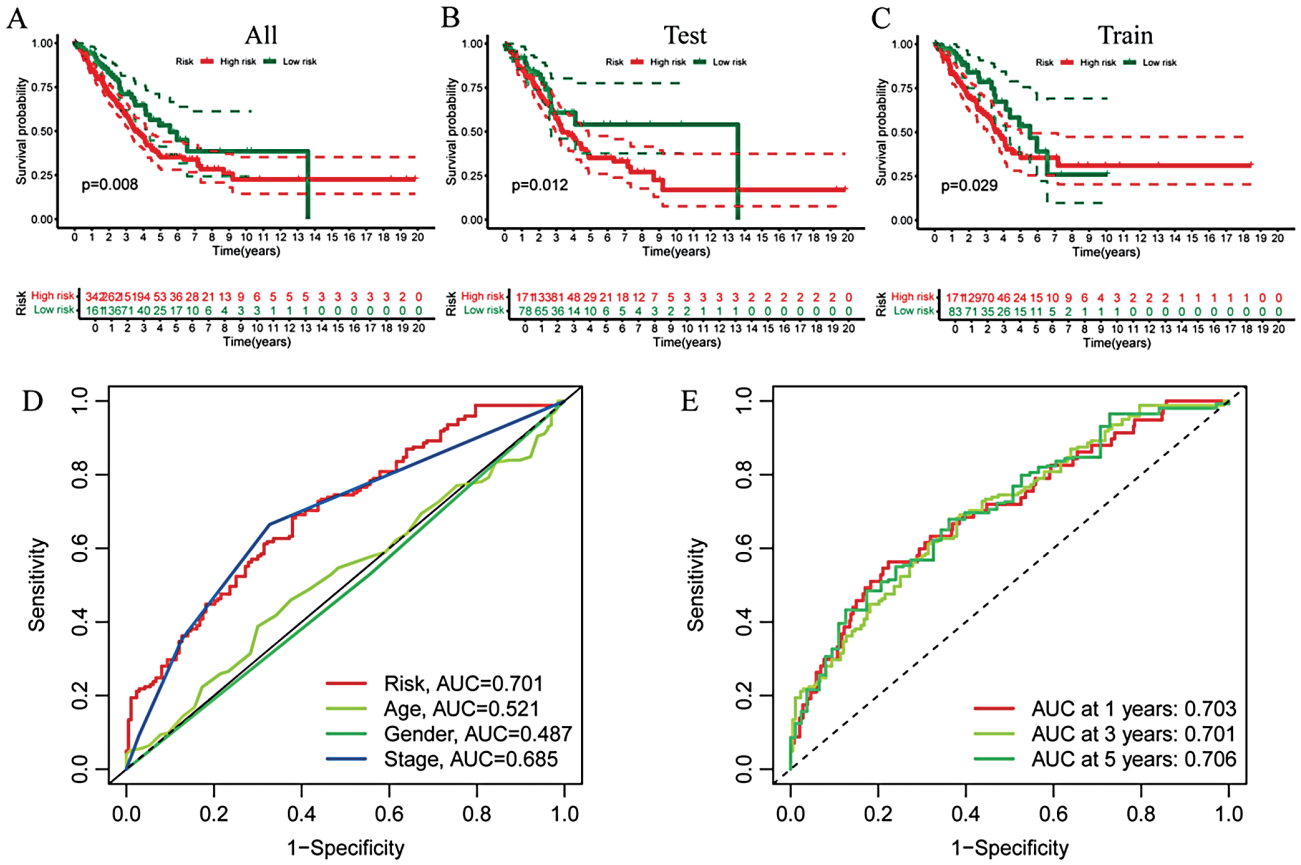

Fig. 3. Validation of the endoplasmic reticulum (ER) stress-related long non-coding RNA (lncRNA) prognostic model. A. Kaplan–Meier survival curves (all); B. Kaplan–Meier survival curves (test); C. Kaplan–Meier survival curves (train); D. Clinically relevant prognostic receiver operating characteristic (ROC); E. Prognostic ROC for risk scores. In the Kaplan–Meier survival curve, red represents the high-risk group and green represents the low-risk group

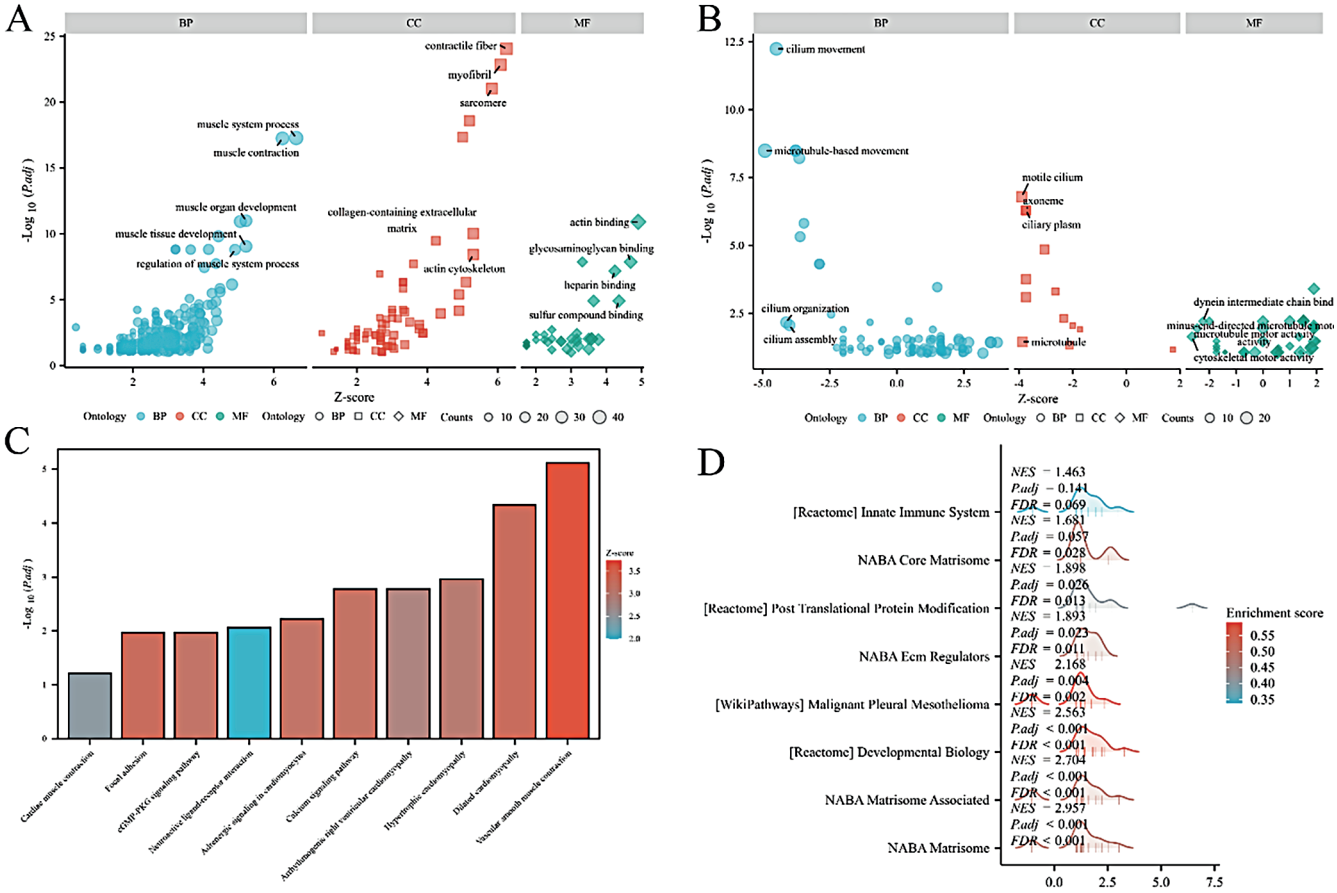

Fig. 4. Functional enrichment analysis for risk score and consensus clustering. A. Gene Ontology (GO) enrichment analysis for the risk score; B. GO enrichment analysis for consensus clustering; C. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for consensus clustering; D. Gene set enrichment analysis (GSEA) enrichment analysis for consensus clustering

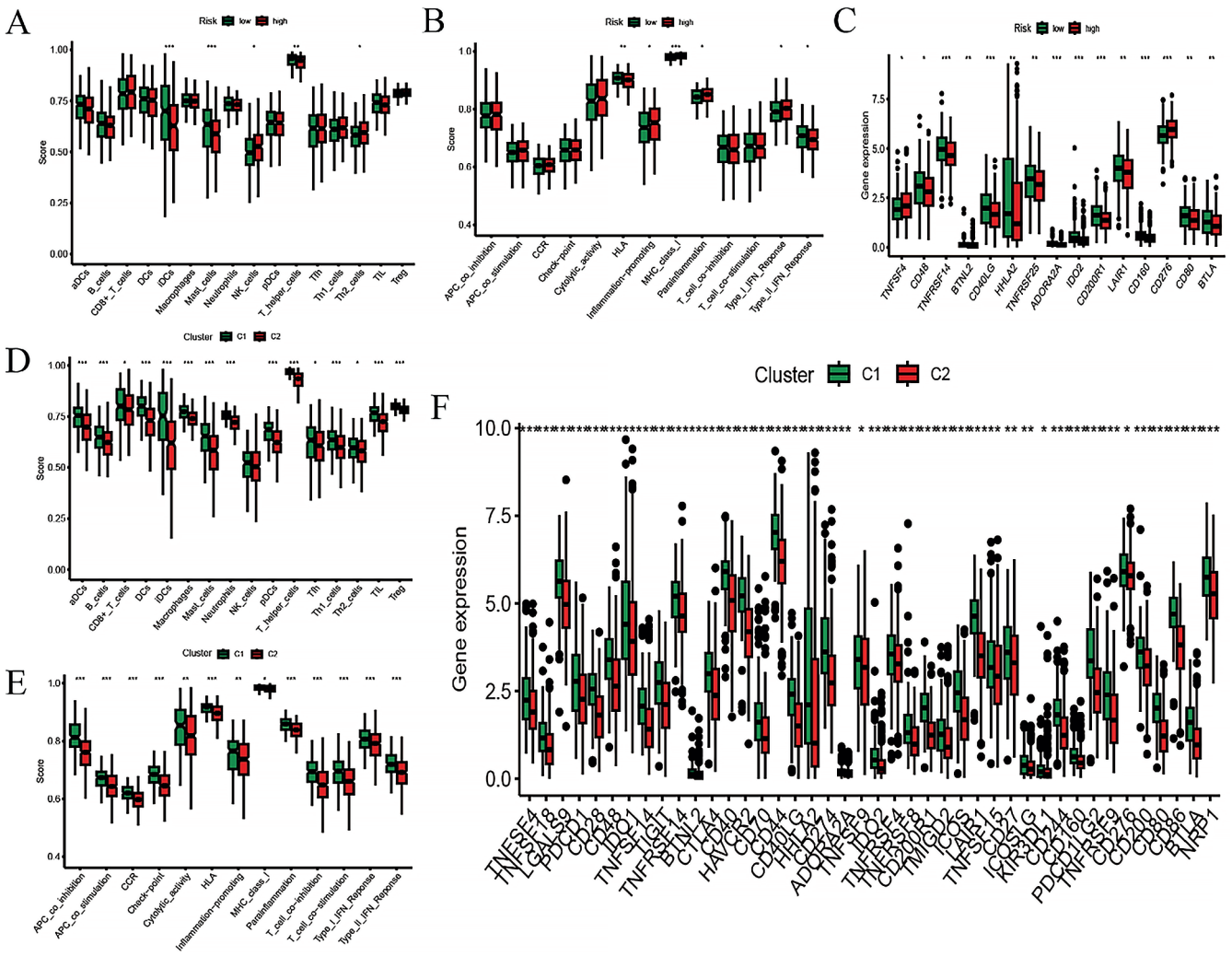

Fig. 5. Immune-related analysis for risk score and consensus clustering. A. Differential analysis of immune cells in risk scoring; B. Differential analysis of immune cell functions in risk scoring; C. Differential analysis of checkpoints in risk scoring; D. Differential analysis of immune cells in consensus clustering; E. Differential analysis of immune cell functions in consensus clustering; F. Differential analysis of checkpoints in consensus clustering. The p-values are represented as *p < 0.05, **p < 0.01, and ***p < 0.001. Green represents C1 or the low-risk group, and red represents C2 or the high-risk group

expression correlation analysis and differential expression analysis, *DNAJC5B* was ultimately selected for subsequent validation. Based on TISCH2, as there was no option for LUAD, we selected NSCLC. *DNAJC5B* was primarily associated with monocytes/macrophages, followed by B cells and natural killer (NK) cells (Fig. 8).

Discussion

The individual variability among patients leads to different responses to treatment, necessitating robust prognostic markers for personalized prediction and therapy.[4,22] Therefore, the construction of a robust prognostic model to predict patient outcomes is urgently needed. Long non-coding RNAs play a significant role in many biological processes, including tumorigenesis, cell differentiation, and metabolism.[23–25] In recent years, these novel non-coding transcripts have garnered widespread attention due to their extensive and complex roles in cancer migration and progression.[26]

Long non-coding RNAs play a key role in cancer development by interacting with ER stress-related proteins, activating downstream signaling pathways, and regulating apoptosis and survival. For example, lncRNA *GAS5* binds to *GRP78*, activates the UPR and the caspase-9 and CHOP signaling pathways, and induces apoptosis in hepatoblastoma HepG2 cells.[27] lncRNA *MEG3* is associated with increased expression of ER stress pathway proteins (e.g., GRP78, PERK, IRE1α, and ATF6), enhances NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells) translocation, and induces apoptosis in cancer cells.[28–30] In addition, lncRNAs are involved in the fine regulation of apoptosis and survival mechanisms and influence the sensitivity of tumor cells to chemotherapeutic agents.[31,32] These mechanisms are complex and diverse, providing potential targets for cancer prognostic assessment and treatment, and further research is needed to support precision cancer therapy in the future.

In the present study, we paid special attention to the role of lncRNAs such as *LINC00460* in ER stress.

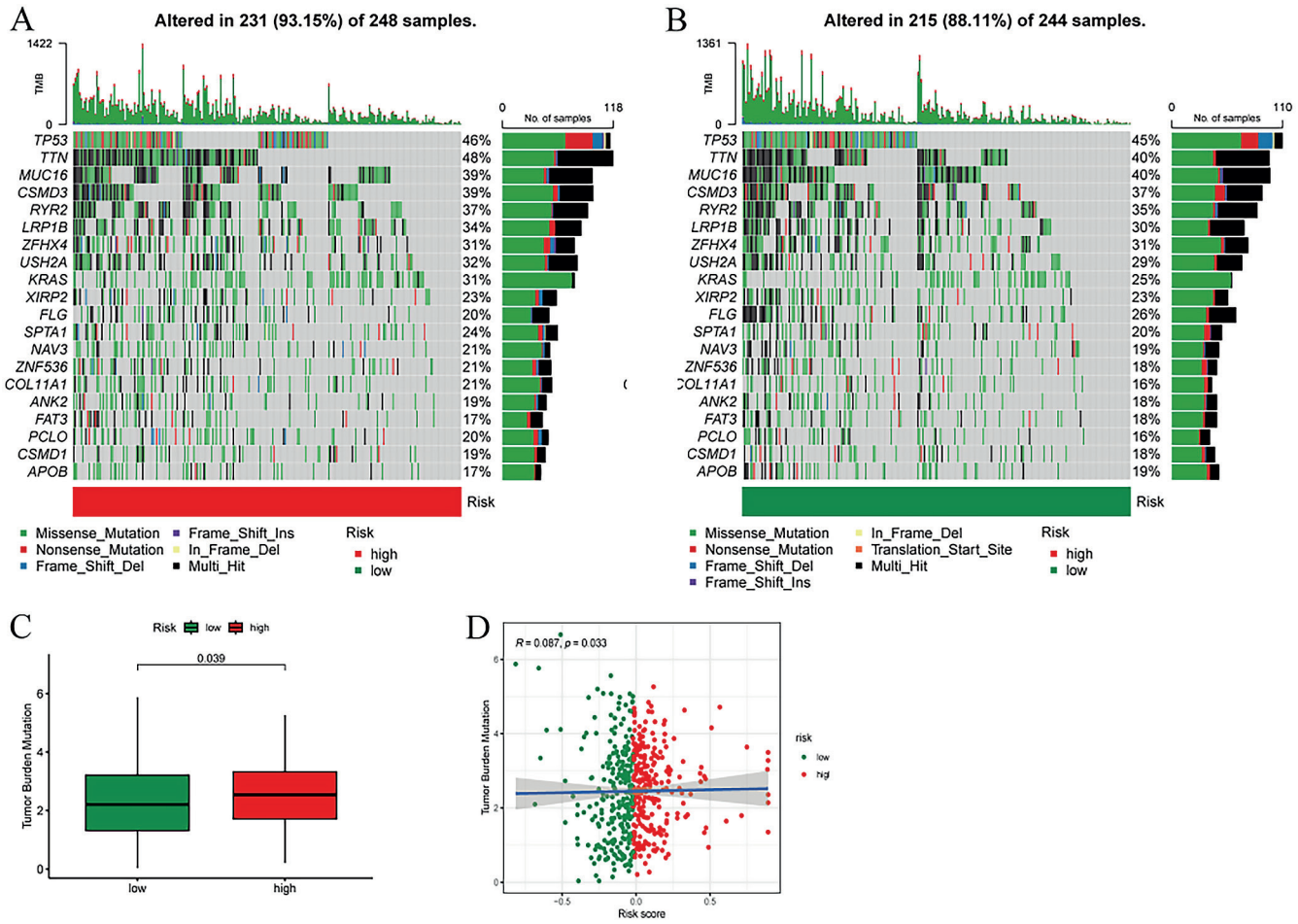

Fig. 6. Exploration of tumor mutation burden based on risk score. A. Mutation landscape of the top 20 genes related to the risk score (high risk); B. Mutation landscape of the top 20 genes related to the risk score (low risk); C. Box plot of tumor mutation burden (TMB) scores for the 2 risk groups; D. Correlation analysis between TMB score and risk score

LINC00460 has been reported to interact with key regulators of ER stress, such as ATF4, which in turn affects UPR signaling. Our data indicate that the expression level of *LINC00460* is closely associated with the prognosis of LUAD patients, suggesting that it may influence tumor progression by regulating the UPR.

We employed the WGCNA method to identify gene modules highly related to the disease from the gene expression data of LUAD patients. Combined with DEGs and univariate Cox analysis, we selected genes significantly associated with prognosis and further optimized model construction using the LASSO Cox regression algorithm. This process not only improved the predictive accuracy of the model but also ensured its robustness through cross-validation.

The risk scoring model and its consensus clustering subgroups demonstrated excellent discrimination ability in both the training and testing sets, clearly distinguishing high- and low-risk patient groups. The predictive accuracy of the model was further validated through Kaplan–Meier survival curves and ROC curve analysis, providing a potential prognostic tool for personalized treatment of LUAD patients. Compared with previous methods that only used the median or average value of the risk score as the threshold to distinguish high and low-risk groups, consensus clustering subgroups (C1 and C2) more accurately combined the risk score with clinical actual subgroups. This combination helps to identify subgroups sensitive to the risk score (C2) and subgroups with unique immune characteristics (C1), providing a more precise stratification basis for clinical treatment. We found that a high risk score was associated with a poorer survival prognosis, which may be related to the high infiltration of immunosuppressive cells in the TME of patients in the high risk score group. This finding provides a potential biomarker for future immunotherapy, suggesting that we can improve patient prognosis by modulating immune cells in the TME.

Furthermore, functional enrichment analysis using the GO and KEGG databases revealed that differential genes were mainly enriched in biological processes related to cell movement and ciliary movement. In contrast, consensus clustering analysis revealed gene enrichment related to muscle contraction and fiber contraction. These findings suggest that ER stress-related lncRNAs may affect the development of LUAD by regulating these key biological processes.

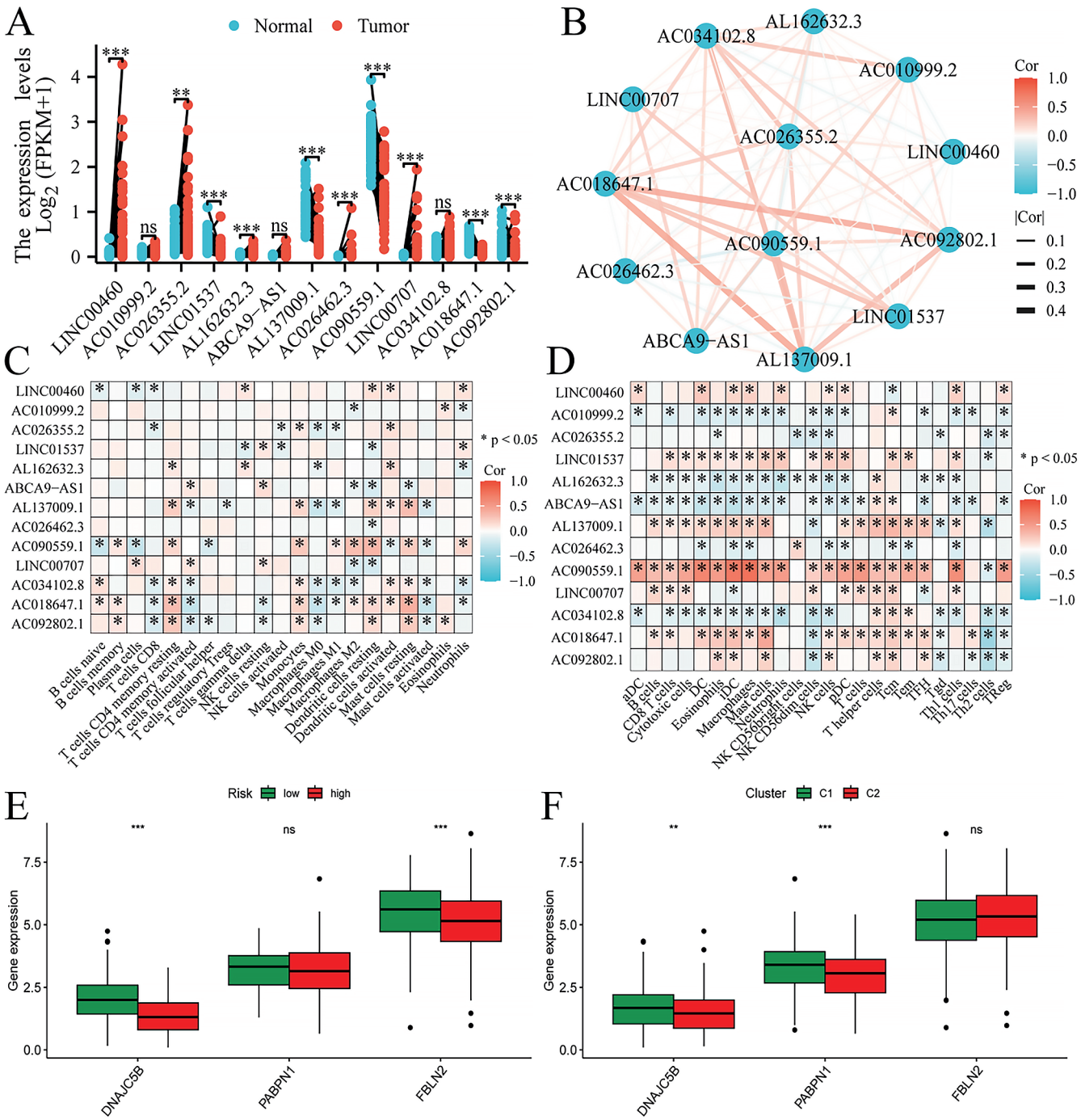

Fig. 7. Clinical and functional exploration of model genes. A. Pairwise differential analysis; B. Correlation network of model genes and target genes; C. Cell and gene correlation (CIBERSORT algorithm); D. Cell and gene correlation (SSGSEA algorithm); E. Differential analysis of core target genes between risk score groups; F. Differential analysis of core target genes between consensus clustering groups. The p-values are represented as *p < 0.05, **p < 0.01, and ***p < 0.001

Muscle contraction and fiber contraction play important roles in tumor biology, as they can regulate vascular contraction and, through the regulation of material exchange between the tumor and the circulatory system, affect tumor metabolism and proliferation on the one hand.[33–35] On the other hand, these processes also participate in the regulation of tumor invasion, metastasis, and immune escape mechanisms.[34,36–38] At the same time, cell movement and ciliary movement directly participate in affecting the invasiveness, metastatic ability, and immune escape mechanisms of tumors.[39,40] It is worth noting that both muscle and fiber contraction and cell and ciliary movement depend on the normal function of the ER to ensure the correct synthesis and execution of protein functions.[41–43] The ER plays a crucial role in maintaining cell movement and tumor biological behavior.

Therefore, the analysis results of this study highlight the multifaceted role of the ER in tumor biology and

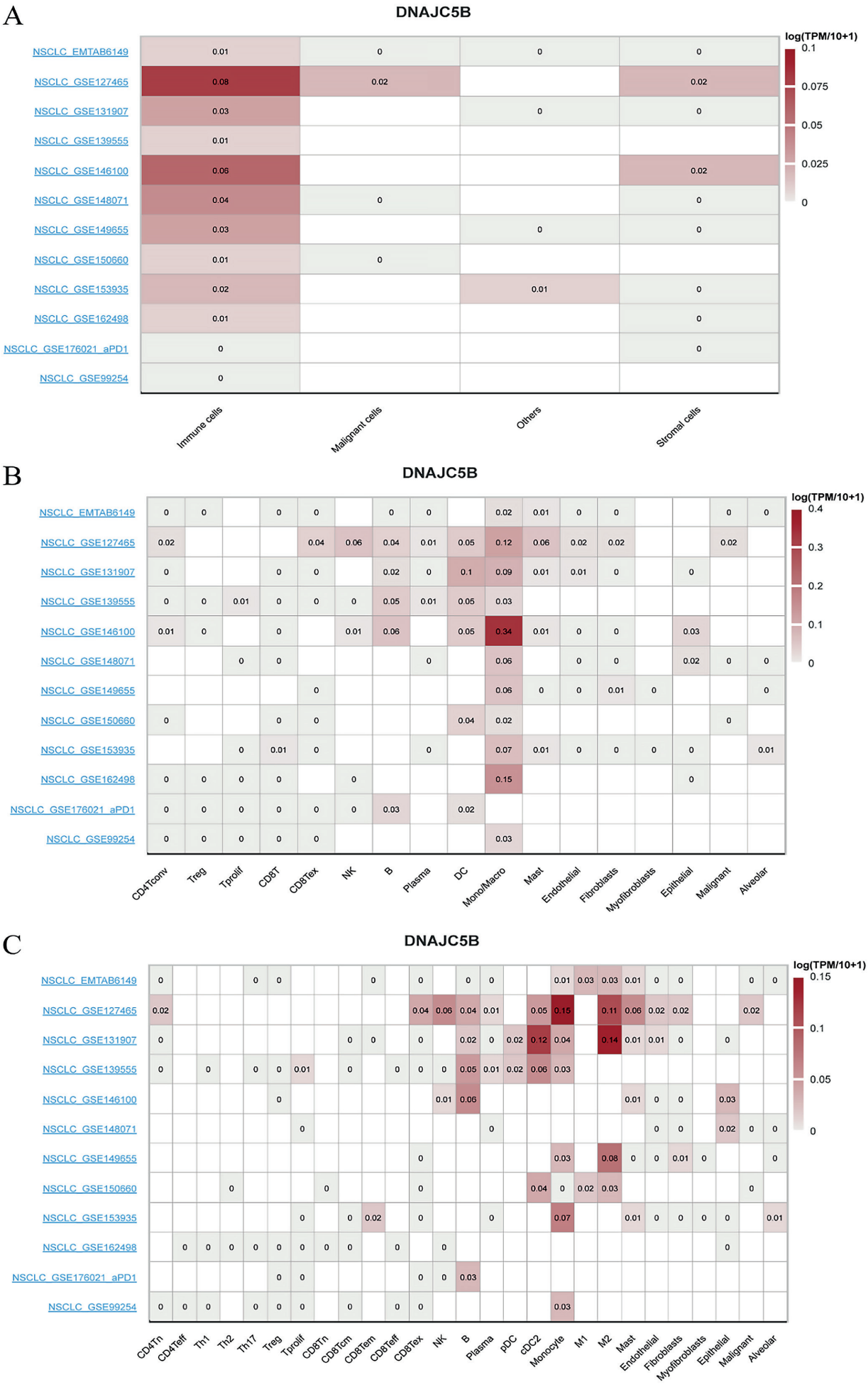

Fig. 8. Single-cell validation. A. Heatmap of malignant tumor groups; B. Heatmap of primary tumor lineage groups; C. Heatmap of secondary tumor lineage groups

provide new perspectives for future research. Especially when developing treatment strategies for LUAD, considering the potential impact of ER stress and related lncRNAs is crucial. These findings may help design new intervention measures to regulate ER stress, thereby affecting tumor development and response to treatment.

The TME, immune system interactions, and TMB are pivotal areas of interest in oncology, significantly influencing tumor genesis and progression.[44–46] Our research demonstrated that the TME score for patients categorized in the high-risk group markedly surpassed that of the low-risk group, suggesting a potential enhanced responsiveness to immunotherapy among high-risk patients. This notion is further corroborated by a detailed examination of immune cell dynamics and their respective roles. Clinically, an elevated TMB score is often linked to enduring clinical benefits and enhanced objective response rates in cancer immunotherapy.[47] Our findings indicate that the TMB score for the high-risk group was considerably higher than for the low-risk group, hinting at a more favorable prognosis and a greater potential for therapeutic advantage from immunotherapy for those with elevated TMB scores.[48]

A combined analysis of risk score grouping and consensus clustering subgroups was performed.[49,50] This finding provides a new perspective for understanding the role of ER stress in the tumor immune microenvironment and may help develop precision treatment strategies for specific immune cell populations.

DNAJC5B (DnaJ heat shock protein family C member 5B), as a member of the heat shock protein family, belongs to the subfamily of DNAJ/Hsp40 protein co-chaperones. These proteins play important roles in cells, including promoting the correct folding of proteins, preventing the aggregation of misfolded proteins, and participating in the degradation and translocation processes of proteins.[51] *DNAJC5B*, with its molecular chaperone characteristics, may be involved in the correct folding of proteins in the ER, preventing the accumulation of abnormal proteins, thereby maintaining protein homeostasis.

When protein folding encounters obstacles, it may trigger ER stress. In this process, *DNAJC5B* may participate in the regulation of the ER stress response through its role in the ER, affecting the cell's response to ER stress. Therefore, *DNAJC5B* may indirectly affect the proliferation, migration, invasion, apoptosis, and colony formation of tumor cells by participating in the ER stress response.[52–54] To verify this hypothesis, we conducted an in-depth exploration of the function of *DNAJC5B* in the experimental part.

In addition, combined with the single-cell sequencing exploration in this study, we found that *DNAJC5B* mainly affects monocytes/macrophages, T cells (NK cells), and B cells. Therefore, we speculate that *DNAJC5B* affects immune cells (monocytes/macrophages, T cells (NK cells), B cells) through the ER stress response, thereby affecting the immune-related treatment of LUAD. Interestingly, the correlation between *DNAJC5B* and immune cells is not related to lung cancer treated with neoadjuvant PD-1 blockade, which provides evidence for our hypothesis.

This study successfully constructed a risk scoring model based on ER stress-related lncRNAs, providing a new perspective and tool for the prognostic evaluation and personalized treatment of LUAD. Through consensus clustering analysis, not only were populations potentially sensitive to treatment identified, but populations with specific immune characteristics were also discovered. In addition, the biological relevance of the core target gene *DNAJC5B* in tumor cell biological behavior was experimentally verified, further confirming the biological relevance of the model.

Limitations of the study

A critical limitation is the absence of external validation cohorts, which may restrict the generalizability of the model. Future studies should prioritize multicenter prospective trials to assess its clinical utility across diverse populations. Additionally, the model's dependency on RNA-seq data necessitates standardization of sequencing platforms for real-world application. The present study's used only data on LUAD patients from the TCGA database, which may be subject to selection bias. In addition, our model has not been validated in an independent cohort, so its generalizability and reliability need to be further assessed. To overcome these limitations, future research will verify the biological function of the model through experimental methods and plans to further explore the clinical application potential of the model in a larger independent patient cohort. Meanwhile, future studies should include validation of our risk scoring model in a larger multicenter cohort of patients and explore its application in guiding individualized treatment. This will help improve the generalizability and reliability of the models, bridge the gap between computational prediction and clinical implementation, and provide more accurate guidance for the clinical treatment of LUAD patients. In addition, we suggest further studies on the regulatory mechanisms of lncRNAs in ER stress and how they affect the immune microenvironment in LUAD.

Conclusions

Our study contributes to the development of new therapeutic strategies to improve outcomes for LUAD patients. This will help improve the model's universality and reliability, providing more accurate guidance for the clinical treatment of LUAD patients. Through these efforts, we hope to translate research results into practical clinical applications, benefiting more LUAD patients.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.16993863. The package contains the following files:

Supplementary File 1. Schoenfeld PH test results.

Data Availability Statement

The datasets supporting the conclusions of this study are publicly available on Figshare at the following link: https://figshare.com/s/0423b4e4fcde57b9dc87.

Consent for publication

All authors have reviewed the manuscript and consent to its publication.

Use of AI and AI-assisted technologies

ORCID iDs

Haiyang Li https://orcid.org/0009-0007-2920-7864
Zhenshan Zhao https://orcid.org/0009-0000-4619-841
Jing Li https://orcid.org/0009-0006-2292-4943
Yao Rong https://orcid.org/0009-0007-1372-1877
Aimin Zheng https://orcid.org/0009-0000-7404-9067
Menghui Hao https://orcid.org/0009-0007-3790-1741

References

1. Cao M, Li H, Sun D, Chen W. Cancer burden of major cancers in China: A need for sustainable actions. *Cancer Commun (Lond)*. 2020;40(5):205–210. doi:10.1002/cac2.12025
2. Ferlay J, Colombet M, Soerjomataram I, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer*. 2018;103:356–387. doi:10.1016/j.ejca.2018.07.005
3. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors. *J Thorac Oncol*. 2015;10(9):1243–1260. doi:10.1097/JTO.0000000000000630
4. Nicholson AG, Tsao MS, Beasley MB, et al. The 2021 WHO Classification of Lung Tumors: Impact of advances since 2015. *J Thorac Oncol*. 2022;17(3):362–387. doi:10.1016/j.jtho.2021.11.003
5. Schwarz DS, Blower MD. The endoplasmic reticulum: Structure, function and response to cellular signaling. *Cell Mol Life Sci*. 2016;73(1):79–94. doi:10.1007/s00018-015-2052-6
6. Balch WE, Morimoto RI, Dillin A, Kelly JW. Adapting proteostasis for disease intervention. *Science*. 2008;319(5865):916–919. doi:10.1126/science.1141448
7. Chen X, Shi C, He M, Xiong S, Xia X. Endoplasmic reticulum stress: Molecular mechanism and therapeutic targets. *Sig Transduct Target Ther*. 2023;8(1):352. doi:10.1038/s41392-023-01570-w
8. Fang C, Weng T, Hu S, et al. IFN-γ-induced ER stress impairs autophagy and triggers apoptosis in lung cancer cells. *OncoImmunology*. 2021;10(1):1962591. doi:10.1080/2162402X.2021.1962591
9. Wan L, Chen Z, Yang J, et al. Identification of endoplasmic reticulum stress-related signature characterizes the tumor microenvironment and predicts prognosis in lung adenocarcinoma. *Sci Rep*. 2023;13(1):19462. doi:10.1038/s41598-023-45690-3
10. Zhu C, Xie Y, Li Q, et al. CPSF6-mediated XBP1 3′UTR shortening attenuates cisplatin-induced ER stress and elevates chemo-resistance in lung adenocarcinoma. *Drug Resist Updat*. 2023;68:100933. doi:10.1016/j.drup.2023.100933
11. Shergalis AG, Hu S, Bankhead A, Neamati N. Role of the ERO1-PDI interaction in oxidative protein folding and disease. *Pharmacol Ther*. 2020;210:107525. doi:10.1016/j.pharmthera.2020.107525
12. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*. 2021;22(2):96–118. doi:10.1038/s41580-020-00315-9
13. Lai X, Zhong J, Zhang A, Zhang B, Zhu T, Liao R. Focus on long non-coding RNA MALAT1: Insights into acute and chronic lung diseases. *Front Genet*. 2022;13:1003964. doi:10.3389/fgene.2022.1003964
14. Mondal P, Natesh J, Kamal MA, Meeran SM. Non-coding RNAs in lung cancer chemoresistance. *Curr Drug Metab*. 2020;20(13):1023–1032. doi:10.2174/1389200221666200106105201
15. Inamura K. Major tumor suppressor and oncogenic non-coding RNAs: Clinical relevance in lung cancer. *Cells*. 2017;6(2):12. doi:10.3390/cells6020012
16. Wei H, Zhang S, Lin X, Fang R, Li L. Differential expression and clinical significance of long non-coding RNAs in the development and progression of lung adenocarcinoma. *Front Oncol*. 2024;14:1411672. doi:10.3389/fonc.2024.1411672
17. Ebrahimi N, Saremi J, Ghanaatian M, et al. The role of endoplasmic reticulum stress in the regulation of long noncoding RNAs in cancer. *J Cell Physiol*. 2022;237(10):3752–3767. doi:10.1002/jcp.30846
18. Li L, Liu B, Wapinski OL, et al. Targeted disruption of hotair leads to homeotic transformation and gene derepression. *Cell Rep*. 2013;5(1):3–12. doi:10.1016/j.celrep.2013.09.003
19. Tripathi V, Ellis JD, Shen Z, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010;39(6):925–938. doi:10.1016/j.molcel.2010.08.011
20. Gonzalez H, Hagerling C, Werb Z. Roles of the immune system in cancer: From tumor initiation to metastatic progression. *Genes Dev*. 2018;32(19–20):1267–1284. doi:10.1101/gad.314617.118
21. Zeng Z, Gao Y, Li J, et al. Violations of proportional hazard assumption in Cox regression model of transcriptomic data in TCGA pan-cancer cohorts. *Comput Struct Biotechnol J*. 2022;20:496–507. doi:10.1016/j.csbj.2022.01.004
22. Ruiz-Cordero R, Devine WP. Targeted therapy and checkpoint immunotherapy in lung cancer. *Surg Pathol Clin*. 2020;13(1):17–33. doi:10.1016/j.path.2019.11.002
23. Tan Y, Lin J, Li T, Li J, Xu R, Ju H. LncRNA-mediated posttranslational modifications and reprogramming of energy metabolism in cancer. *Cancer Commun (Lond)*. 2021;41(2):109–120. doi:10.1002/cac2.12108
24. McCabe EM, Rasmussen TP. lncRNA involvement in cancer stem cell function and epithelial–mesenchymal transitions. *Semin Cancer Biol*. 2021;75:38–48. doi:10.1016/j.semcancer.2020.12.012
25. Yan H, Bu P. Non-coding RNA in cancer. *Essays Biochem*. 2021;65(4):625–639. doi:10.1042/EBC20200032
26. Winkle M, El-Daly SM, Fabbri M, Calin GA. Noncoding RNA therapeutics: Challenges and potential solutions. *Nat Rev Drug Discov*. 2021;20(8):629–651. doi:10.1038/s41573-021-00219-z
27. Zhang W, Zhan H, Li M, Wu G, Liu Z, Wu L. Long noncoding RNA Gas5 induces cell apoptosis and inhibits tumor growth via activating the CHOP-dependent endoplasmic reticulum stress pathway in human hepatoblastoma HepG2 cells. *J Cell Biochem*. 2022;123(2):231–247. doi:10.1002/jcb.30159
28. Zhang Y, Wu J, Jing H, Huang G, Sun Z, Xu S. Long noncoding RNA *MEG3* inhibits breast cancer growth via upregulating endoplasmic reticulum stress and activating NF-κB and p53. *J Cell Biochem*. 2019;120(4):6789–6797. doi:10.1002/jcb.27982
29. Huang ZL, Chen RP, Zhou XT, et al. Long non-coding RNA MEG3 induces cell apoptosis in esophageal cancer through endoplasmic reticulum stress. *Oncol Rep*. 2017;37(5):3093–3099. doi:10.3892/or.2017.5568
30. Chen RP, Huang ZL, Liu LX, et al. Involvement of endoplasmic reticulum stress and p53 in lncRNA MEG3-induced human hepatoma HepG2 cell apoptosis. *Oncol Rep*. 2016;36(3):1649–1657. doi:10.3892/or.2016.4919
31. Ding Z, Kang J, Yang Y. Long non-coding RNA CASC2 enhances irradiation-induced endoplasmic reticulum stress in NSCLC cells through PERK signaling. *3 Biotech*. 2020;10(10):449. doi:10.1007/s13205-020-02443-7
32. Tsuchiya H, Shinonaga R, Sakaguchi H, Kitagawa Y, Yoshida K. NEAT1–SOD2 axis confers sorafenib and lenvatinib resistance by activating AKT in liver cancer cell lines. *Curr Issues Mol Biol*. 2023;45(2):1073–1085. doi:10.3390/cimb45020071

33. Kuczynski EA, Vermeulen PB, Pezzella F, Kerbel RS, Reynolds AR. Vessel co-option in cancer. *Nat Rev Clin Oncol*. 2019;16(8):469–493. doi:10.1038/s41571-019-0181-9
34. Rajabi S, Dehghan MH, Dastmalchi R, Jalali Mashayekhi F, Salami S, Hedayati M. The roles and role-players in thyroid cancer angiogenesis. *Endocr J*. 2019;66(4):277–293. doi:10.1507/endocrj.EJ18-0537
35. Pieterse Z, Sinha D, Kaur P. Pericytes in metastasis. In: Birbrair A, ed. *Pericyte Biology in Disease*. Vol. 1147. Advances in Experimental Medicine and Biology. Cham, Switzerland: Springer International Publishing; 2019:125–135. doi:10.1007/978-3-030-16908-4_5
36. Kuczynski EA, Reynolds AR. Vessel co-option and resistance to anti-angiogenic therapy. *Angiogenesis*. 2020;23(1):55–74. doi:10.1007/s10456-019-09698-6
37. Lugassy C, Vermeulen PB, Ribatti D, Pezzella F, Barnhill RL. Vessel co-option and angiotropic extravascular migratory metastasis: A continuum of tumour growth and spread? *Br J Cancer*. 2022;126(7):973–980. doi:10.1038/s41416-021-01686-2
38. Li S, Zhang Q, Hong Y. Tumor vessel normalization: A window to enhancing cancer immunotherapy. *Technol Cancer Res Treat*. 2020;19:1533033820980116. doi:10.1177/1533033820980116
39. Zanotelli MR, Zhang J, Reinhart-King CA. Mechanoresponsive metabolism in cancer cell migration and metastasis. *Cell Metab*. 2021;33(7):1307–1321. doi:10.1016/j.cmet.2021.04.002
40. Marcadis AR, Kao E, Wang Q, et al. Rapid cancer cell perineural invasion utilizes amoeboid migration. *Proc Natl Acad Sci U S A*. 2023;120(17):e2210735120. doi:10.1073/pnas.2210735120
41. Weigel AV, Chang CL, Shtengel G, et al. ER-to-Golgi protein delivery through an interwoven, tubular network extending from ER. *Cell*. 2021;184(9):2412–2429.e16. doi:10.1016/j.cell.2021.03.035
42. Battson ML, Lee DM, Gentile CL. Endoplasmic reticulum stress and the development of endothelial dysfunction. *Am J Physiol Heart Circ Physiol*. 2017;312(3):H355–H367. doi:10.1152/ajpheart.00437.2016
43. Nieblas B, Pérez-Treviño P, García N. Role of mitochondria-associated endoplasmic reticulum membranes in insulin sensitivity, energy metabolism, and contraction of skeletal muscle. *Front Mol Biosci*. 2022;9:959844. doi:10.3389/fmolb.2022.959844
44. Xiao Y, Yu D. Tumor microenvironment as a therapeutic target in cancer. *Pharmacol Ther*. 2021;221:107753. doi:10.1016/j.pharmthera.2020.107753
45. Kiely M, Lord B, Ambs S. Immune response and inflammation in cancer health disparities. *Trends Cancer*. 2022;8(4):316–327. doi:10.1016/j.trecan.2021.11.010
46. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet*. 2019;51(2):202–206. doi:10.1038/s41588-018-0312-8
47. Allgäuer M, Budczies J, Christopoulos P, et al. Implementing tumor mutational burden (TMB) analysis in routine diagnostics: A primer for molecular pathologists and clinicians. *Transl Lung Cancer Res*. 2018;7(5):703–715. doi:10.21037/tlcr.2018.08.14
48. Zhang Y, Wang L, Li R, Liu B. The emerging development of tumor mutational burden in patients with NSCLC. *Future Oncol*. 2020;16(9):469–481. doi:10.2217/fon-2019-0650
49. Li F, Ma J, Yan C, Qi Y. ER stress-related mRNA-lncRNA co-expression gene signature predicts the prognosis and immune implications of esophageal cancer. *Am J Transl Res*. 2022;14(11):8064–8084. PMID:36505280. PMCID:PMC9730056.
50. Yi J, Wang L, Hu G, et al. CircPVT1 promotes ER-positive breast tumorigenesis and drug resistance by targeting *ESR1* and MAVS. *EMBO J*. 2023;42(10):e112408. doi:10.15252/embj.2022112408
51. Braga ACS, Carneiro BM, Batista MN, Akinaga MM, Bittar C, Rahal P. Heat shock proteins HSPB8 and DNAJC5B have HCV antiviral activity. *PLoS One*. 2017;12(11):e0188467. doi:10.1371/journal.pone.0188467
52. Sun H, Cai X, Zhou H, et al. The protein–protein interaction network and clinical significance of heat-shock proteins in esophageal squamous cell carcinoma. *Amino Acids*. 2018;50(6):685–697. doi:10.1007/s00726-018-2569-8
53. Zhou MH, Wang XK. Microenvironment-related prognostic genes in esophageal cancer. *Transl Cancer Res*. 2020;9(12):7531–7539. doi:10.21037/tcr-20-2288
54. Mirzaei MR, Asadi M, Mowla SJ, Hassanshahi G, Ahmadi Z. Down-regulation of *HSP40* gene family following OCT4B1 suppression in human tumor cell lines. *Iran J Basic Med Sci*. 2016;19(2):187–193. PMID:27081464. PMCID:PMC4818367.

Inflammatory cytokines and metabolic pathways in osteomyelitis: Mendelian randomization insights and experimental validation

Xingyu Chen[1,A–D], Ruiqing Mo[1,E], Sijie Yang[2,E], Peilin Zhou[1,E,F], Hua Qikai[1,3,F]

[1] Department of Bone and Joint Surgery, Guangxi Diabetic Foot Salvage Engineering Research Center for Regenerative Medicine, The First Affiliated Hospital of Guangxi Medical University, Nanning, China
[2] Department of Plastic and Reconstructive Surgery, The People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, China
[3] Collaborative Innovation Centre of Regenerative Medicine and Medical BioResource Development and Application Co-constructed by the Province and Ministry, Guangxi Medical University, Nanning, 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(5):863–876

Address for correspondence
Hua Qikai
E-mail: hqk100@yeah.net

Funding sources
This study was supported by the National Natural Science Foundation of China (grant No. 82260448), the Guangxi Natural Science Foundation (grant No. 2023JJD140126), the Guangxi Key Research and Development Plan (grant No. 2021AB11027), the Key Research and Development Plan of Qingxiu District, Nanning (grant No. 2020053), and the Clinical Research Climbing Plan of the First Affiliated Hospital of Guangxi Medical University (grant No. YYZS2020010).

Conflict of interest
None declared

Received on July 9, 2025
Reviewed on September 8, 2025
Accepted on September 20, 2025

Published online on May 27, 2026

Cite as
Chen X, Mo R, Yang S, Zhou P, Qikai H. Inflammatory cytokines and metabolic pathways in osteomyelitis: Mendelian randomization insights and experimental validation. *Adv Clin Exp Med*. 2026;35(5):863–876. doi:10.17219/acem/211133

DOI
10.17219/acem/211133

Abstract

Background. Osteomyelitis is a challenging orthopedic condition characterized by bone inflammation, often resulting from infection. The roles of inflammatory cytokines (ICs) and metabolites in its pathogenesis are not fully understood.

Objectives. This study aimed to explore the potential causal relationships among ICs, metabolites, and osteomyelitis using Mendelian randomization (MR) analysis and in vitro experiments.

Materials and methods. A 2-sample MR analysis (FinnGen: 1,881 cases and 391,037 controls) was performed to screen 91 cytokines and more than 1,400 metabolites. In vitro experiments using MC3T3-E1 cells treated with staphylococcal protein A (SPA) were conducted to evaluate the effects of p-cresol sulfate (PCS), a circulating metabolite identified through MR analysis, on cell proliferation, osteogenic differentiation, and inflammation.

Results. Mendelian randomization analysis identified significant associations between several ICs and osteomyelitis risk. Elevated levels of CUB domain-containing protein 1 (*CDCP1*; odds ratio (OR) = 1.19, 95% confidence interval (95% CI): 1.03–1.38) and thymic stromal lymphopoietin (*TSLP*; OR = 1.26, 95% CI: 1.05–1.51) were associated with increased risk, whereas higher levels of adenosine deaminase (*ADA*; OR = 0.85, 95% CI: 0.74–0.98) and interleukin-5 (*IL-5*; OR = 0.79, 95% CI: 0.65–0.96) were associated with reduced risk. Metabolites such as PCS (OR = 0.74, 95% CI: 0.59–0.93) were identified as protective, whereas others, such as beta-cryptoxanthin (β-CX; OR = 1.33, 95% CI: 1.07–1.66), were associated with increased risk. Mediation analysis further suggested that several metabolites significantly mediated the indirect effects of cytokines on osteomyelitis risk. In vitro experiments demonstrated that PCS enhanced osteogenic potential, reduced intracellular reactive oxygen species (ROS) production, and lowered IC levels.

Conclusions. These findings provide insight into the associations among ICs, metabolites, and osteomyelitis, suggesting potential therapeutic targets for reducing disease severity. They also highlight the complex interplay among these factors in osteomyelitis pathogenesis. Further studies are needed to clarify the mechanisms through which ICs and metabolites influence osteomyelitis, particularly through regulation of inflammatory responses and oxidative stress.

Key words: osteomyelitis, inflammatory cytokines, metabolomics, Mendelian randomization, osteoblasts

Highlights

- Mendelian randomization (MR) analysis identifies key inflammatory cytokines and metabolites linked to osteomyelitis risk and disease progression.
- Protective metabolite p-cresol sulfate (PCS) enhances osteogenic differentiation and reduces oxidative stress in osteomyelitis models.
- Cytokines including *CDCP1*, *TSLP*, *ADA*, and *IL-5* emerge as potential biomarkers and therapeutic targets for osteomyelitis.
- Integrated genetic and in vitro evidence reveals complex interactions between inflammation, metabolism, and bone infection pathogenesis.

Background

Osteomyelitis is an inflammatory bone condition that can cause damage to both bone and adjacent tissues.[1] This disorder may occur at any age and affect any bone, often resulting from the local spread of infection from a contiguous source, such as trauma, bone surgery, or joint replacement.[2] Owing to its complex and chronic nature, osteomyelitis is considered a refractory condition in orthopedics and poses a significant challenge in orthopedic surgical practice.[3,4] Even with optimal medical management, approx. 40% of treatments for chronic or recurrent osteomyelitis fail, partly due to persistent infection and bacterial persistence.[5]

Recent studies have emphasized the key roles of cytokines such as interleukin-1 beta (*IL-1β*) and interleukin-6 (*IL-6*) in osteomyelitis pathophysiology, particularly in osteoblast–osteoclast communication.[6,7] The importance of the precise temporal and spatial expression of these cytokines in maintaining the balance between inflammation and immunoregulation is increasingly recognized.[8] Gut-derived indole derivatives and short-chain fatty acids have recently been shown to modulate both systemic inflammation and bone turnover.[9,10] In high-income countries, the annual incidence of osteomyelitis is approx. 8–10 per 100,000 person-years, with direct medical costs exceeding USD 1 billion annually.[11,12]

Mendelian randomization (MR) uses genetic variation as a natural instrument to investigate potential causal relationships between modifiable risk factors and health outcomes in observational datasets.[13] Because genetic variants are fixed at birth, MR is less susceptible to confounding than conventional observational studies. Given that the associations among inflammatory cytokines (ICs), metabolites, and osteomyelitis have not yet been examined using genetic approaches, we employed a 2-sample MR approach to assess the potential causal relationships between genetically predicted blood concentrations of ICs and metabolites and the risk of osteomyelitis.

While MR is a powerful tool for exploring potential causal relationships and identifying ICs associated with osteomyelitis risk, it is limited in its ability to elucidate the specific cellular mechanisms through which these ICs exert their effects. Staphylococcal protein A (SPA), a key virulence factor of *Staphylococcus aureus* that impairs osteoblast function, is widely used to model infection-induced bone damage. It binds to osteoblasts, inhibits their proliferation, induces apoptosis, and impairs mineralization.[14]

Objectives

The primary aim of this study was to screen for potential causal effects of 91 cytokines and more than 1,400 blood metabolites on osteomyelitis risk using a 2-sample MR approach. Secondary aims were to: 1) perform mediation analysis to quantify indirect pathways; 2) conduct reverse-direction MR analysis; and 3) validate the effects of PCS in vitro using SPA-treated MC3T3-E1 cells. We hypothesized that p-cresol sulfate (PCS) would attenuate SPA-induced oxidative stress and IC release in osteoblasts.

Materials and methods

Study design

In the 1st stage, the inverse-variance weighting (IVW) method was used as the primary MR analysis, alongside MR-Egger, weighted median, simple mode, and weighted mode methods to improve robustness and address potential pleiotropy.[15,16] Mediation was assessed using a 2-step MR framework. In step 1, 2-sample MR was used to estimate the total effect of each cytokine on osteomyelitis and the direct effect of each metabolite on osteomyelitis. In step 2, 2-sample MR was used to estimate the effect of each cytokine on its candidate mediator (metabolite). The proportion mediated (PM) was calculated as $PM = (\alpha \times \beta)/\gamma$, where α represents the effect of the cytokine on the metabolite, β represents the effect of the metabolite on osteomyelitis, and γ represents the total effect of the cytokine on osteomyelitis. Standard errors (SEs) for PM were estimated using the delta method. Inflammatory cytokine data were obtained from the largest available protein quantitative trait

loci (pQTL) study (GWAS Catalog IDs GCST90274758–GCST90274848), which quantified 91 plasma proteins using the Olink Target Inflammation platform in 14,824 individuals of European ancestry (Fig. 1).[17] Plasma metabolite genome-wide association studies (GWAS) data were derived from the Canadian Longitudinal Study on Aging (n = 8,299 participants of European ancestry; Metabolon HD4 platform; GWAS Catalog IDs GCST90199621–GCST9021020). Osteomyelitis summary statistics (1,881 cases and 391,037 controls) were obtained from the FinnGen R10 release (Finnish population). Mean F-statistics were 31 (range: 10–120) for cytokines and 29 (range: 10–105) for metabolites, indicating robust instrument strength.

Genetic instrumental variables selection

Single nucleotide polymorphisms (SNPs) were selected as instrumental variables (IVs) for IC and metabolite traits using a p-value threshold of $<1 \times 10^{-5}$.[18] For osteomyelitis, a more stringent threshold of $p < 5 \times 10^{-8}$ was applied. Genetic variants were clumped using an R^2 threshold of <0.001 within a 10,000-kb window. Single nucleotide polymorphisms were further filtered based on the F-statistic, calculated as the ratio of the squared effect size (β^2) to the squared standard error (SE^2), with a threshold of >10. Potential confounders and related traits (e.g., age, sex, ethnicity, and comorbid conditions) were identified using the PhenoScanner V2 database (http://www.phenoscanner.medschl.cam.ac.uk/).

Allele harmonization was performed using the TwoSampleMR package (action = 2), which automatically aligned ambiguous (palindromic) SNPs to the positive strand and ensured concordant effect alleles across the exposure and outcome datasets. Linkage disequilibrium clumping was performed using the 1000 Genomes European reference panel (https://github.com/MRCIEU/gwasvcf) with a 10,000-kb window and an r^2 threshold of <0.001. MR-PRESSO (Mendelian Randomization Pleiotropy RESidual Sum and Outlier) was used to assess horizontal pleiotropy: the global test evaluated overall pleiotropy, whereas the outlier test identified influential SNPs. Outliers with distortion $p < 0.05$ were removed, and MR estimates were recalculated using the cleaned dataset.

Allele harmonization and strand alignment were performed using the TwoSampleMR package (v. 0.5.6) following the protocol described by Hemani et al.[19] To contextualize our findings within recent large-scale studies, we note the comprehensive MR overview by Wang et al.[20] and the proteome-wide investigation by Sun et al.[17] Our study extends these efforts by focusing specifically on osteomyelitis.

Cell culture and amplification

The MC3T3-E1 cell line was obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). Cells between passages 3 and 6 were used for all experiments. All experiments were performed with at least 3 independent biological replicates (n = 3 per group). Cells were randomly allocated to treatment groups, and the operator was blinded to treatment identity during image acquisition and quantification. To maintain the cells, they were cultured in phenol-red-free Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated fetal bovine serum (FBS-HI), 100 U/mL penicillin, and 100 μg/mL streptomycin. Cells were incubated at 37°C in a humidified atmosphere containing 5% CO_2.

Osteomyelitis in vitro model

A 50 μg/mL SPA solution was added to the culture plates and incubated at 37°C, with the concentration selected based on a previous study.[21] After incubation, unbound SPA was removed by aspiration. Cells (5×10^5 cells/mL) were then seeded into each well. The co-culture was maintained for 4 h.

Cell proliferation assay

The Cell Counting Kit-8 (CCK-8) assay (MedChemExpress, Monmouth Junction, USA) was used to evaluate cell viability. MC3T3-E1 cells were seeded in 96-well plates at a density of 4×10^3 cells per well. Following 2-day exposure to different concentrations of PCS (0, 10, 50, and 100 μg/mL), 10 μL of CCK-8 reagent was added to each well, and the plates were incubated at 37°C for 2 h. Absorbance at 450 nm was measured using a spectrophotometer (Molecular Devices, San Jose, USA) to assess cell viability. Based on the preceding CCK-8 dose–response curve (0–100 μg/mL), 50 μg/mL was selected as the non-cytotoxic concentration that maximally restored osteoblast function and was therefore used in all subsequent alkaline phosphatase (ALP), alizarin red staining (ARS), and reactive oxygen species (ROS) assays.

Alkaline phosphatase staining

MC3T3-E1 cells were cultured in osteogenic induction medium for 7 days. For differentiation assays, cells were seeded at a density of 5×10^4 cells per well. On day 7, cells were stained using the Cell Alkaline Phosphatase Staining Kit (Solarbio, Beijing, China) according to the manufacturer's instructions. Following staining, the cells were observed under an optical microscope (Leica DMRB Microscope; Leica Microsystems GmbH, Wetzlar, Germany), and representative images were captured.

Alizarin red staining

MC3T3-E1 cells were cultured in osteogenic induction medium for 21 days. After the induction period, the medium was removed, and the cells were stained

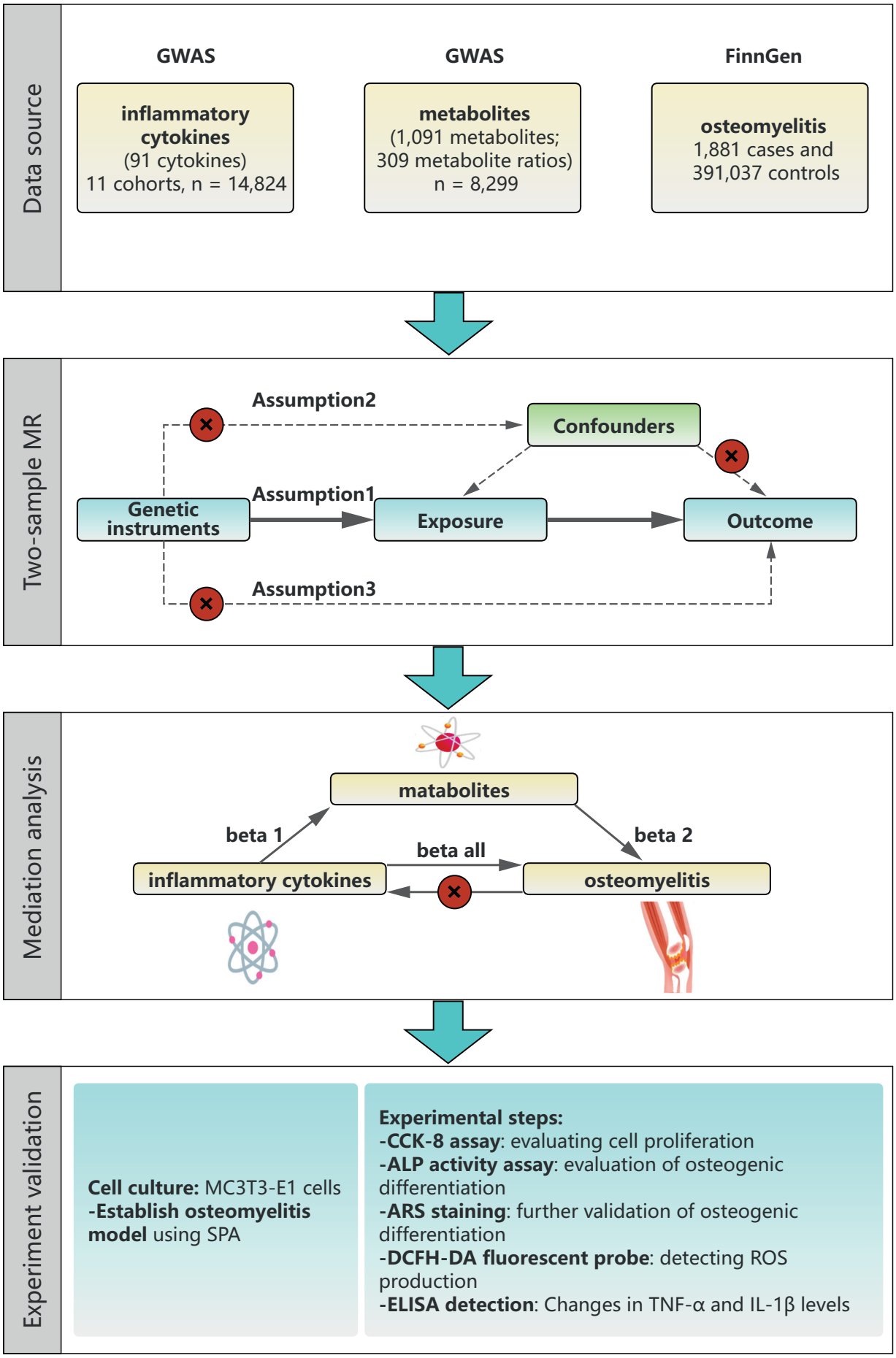

Fig. 1. Study flowchart

GWAS – gene-wide association study; CCK-8 assay – Cell Counting Kit-8 assay; ROS – reactive oxygen species; ELISA – enzyme-linked immunosorbent assay; MR – Mendelian randomization; ALP – alkaline phosphatase; ARS – Alizarin Red S; SPA – staphylococcal protein A; DCFH-DA – 2',7'-dichlorodihydrofluorescein diacetate ;TNF-α – tumor necrosis factor alpha; IL-1β – interleukin-1 beta.

using the Cell Alizarin Red S Staining Kit (Solarbio) according to the manufacturer's instructions. The stained cells were then examined under a microscope (Leica Microsystems GmbH), and representative images were captured.

Detection of intracellular O_2 production

The cell-permeable 2′,7′-dichlorodihydrofluorescein diacetate (DCFH-DA; MedChemExpress) probe was used to measure intracellular ROS levels. For ROS assays, cells were seeded at a density of 2×10^4 cells per well. SPA-treated MC3T3-E1 cells were co-cultured with PCS for 3 days. After rinsing with phosphate-buffered saline (PBS), the cells were incubated with 10 μM DCFH-DA in the dark for 30 min, then washed twice with PBS. Fluorescence images were acquired using a fluorescence microscope (Nikon Eclipse 80i; Nikon Corp., Tokyo, Japan).

ELISA

The culture supernatants from SPA-treated MC3T3-E1 cells were collected, and changes in tumor necrosis factor alpha (TNF-α) and IL-1β levels were measured using commercially available Yuanju Bio, Shanghai, China) enzyme-linked immunosorbent assay (ELISA) kits according to the manufacturer's instructions. Each measurement was performed in triplicate. Concentrations were calculated based on standard calibration curves, and mean values were subsequently determined.

Statistical analyses

All statistical analyses were performed using R v. 4.3.3 (R Foundation for Statistical Computing, Vienna, Austria) with the TwoSampleMR package (https://github.com/MRCIEU/TwoSampleMR). To account for the multiple-testing burden arising from approx. 1,500 exposures, we applied a 2-stage filtering strategy. First, associations with $p < 0.05$ based on the IVW estimator were retained. Second, these p-values were corrected using the Benjamini–Hochberg false discovery rate (FDR) procedure across all independent tests (91 cytokines and 1,400 metabolites analyzed separately). Only cytokines or metabolites with FDR-adjusted $p < 0.05$ were considered statistically significant.

For cellular experiments, all data were analyzed using GraphPad Prism v. 10 (GraphPad Software, San Diego, USA). CCK-8 proliferation data were analyzed using 2-way analysis of variance (ANOVA; factors: treatment and time), followed by Dunnett's multiple-comparisons test to obtain adjusted p-values. All other cellular assays (ALP activity, ARS quantification, ROS fluorescence intensity, and ELISA) were analyzed using unpaired 2-tailed t-tests after confirmation of normal distribution with the Shapiro–Wilk test. Count data were analyzed using the χ^2 test. Results are presented as mean ± standard deviation (SD) from at least 3 independent biological replicates (passages 3–6). Statistical significance was set at p (or adjusted p) <0.05.

Ethics approval and biosafety

The use of the MC3T3-E1 cell line was approved by the Institutional Animal Care and Use Committee (IACUC) of Shanghai Jiao Tong University School of Medicine (China). As this study was conducted entirely in vitro using a commercially available cell line and did not involve live animals, an animal ethics approval number is not applicable. All procedures involving SPA were conducted under Biosafety Level 2 (BSL-2) containment in accordance with institutional guidelines.

This study relied on publicly available summary statistics from previously published studies and research consortia. Ethical approval had been obtained by the respective institutional review boards for each original study, and all participants had provided informed consent. As this study did not involve individual-level data, no additional ethical approval was required.

Results

Inflammatory cytokines as dual modulators in osteomyelitis risk

A total of 33,384, 1,397, and 229 SNPs were selected as IVs for 1,400 metabolites, 91 ICs, and osteomyelitis, respectively (Supplementary Tables 1–3). Mendelian randomization analysis revealed significant associations between IC levels and osteomyelitis risk. The primary analysis, conducted using the IVW method (Fig. 2), indicated that elevated levels of *CDCP1* were associated with an increased risk of osteomyelitis (odds ratio (OR) = 1.19, $p < 0.05$). Similarly, higher levels of thymic stromal lymphopoietin (*TSLP*) were associated with increased osteomyelitis risk (OR = 1.26, $p < 0.05$). Conversely, certain ICs, including adenosine deaminase (*ADA*) (OR = 0.85, $p < 0.05$) and *IL-5* (OR = 0.79, $p < 0.05$), showed protective associations with osteomyelitis. These findings highlight the potentially diverse roles of ICs in osteomyelitis pathophysiology and identify promising candidates for further investigation as potential therapeutic targets. MR-PRESSO and MR-Egger analyses were used to assess pleiotropy (Supplementary Table 4). Heterogeneity testing using Cochran's Q (Supplementary Table 4) yielded $p > 0.05$ for all 3 cytokines across both IVW and MR-Egger models, indicating low heterogeneity. Leave-one-out analysis (Supplementary Fig. 1,2) demonstrated that no single SNP had a disproportionate impact on the causal estimates, supporting the robustness of the results. Potential horizontal pleiotropy was assessed using MR-PRESSO and, where appropriate, multivariable MR (MVMR) to adjust for correlated exposures.[22,23] We then conducted reverse MR analysis to evaluate the effect of osteomyelitis on these

Trails	method	nsnp	pval	OR(95%CI)
Adenosine Deaminase levels	MR Egger	5	0.3198	0.9012 (0.7591–1.0697)
	Weighted median	5	0.0589	0.8689 (0.7510–1.0053)
	Inverse variance weighted	5	<0.05	0.8534 (0.7440–0.9788)
	Simple mode	5	0.4191	0.8671 (0.6357–1.1829)
	Weighted mode	5	0.1214	0.8697 (0.7564–0.9999)
beta–nerve growth factor levels	MR Egger	35	<0.05	0.5458 (0.3318–0.8976)
	Weighted median	35	0.0934	0.7768 (0.5783–1.0435)
	Inverse variance weighted	35	<0.05	0.8007 (0.6463–0.9919)
	Simple mode	35	0.2741	0.7225 (0.4073–1.2816)
	Weighted mode	35	0.2001	0.7295 (0.4545–1.1708)
Eotaxin levels	MR Egger	27	0.0979	0.7489 (0.5386–1.0413)
	Weighted median	27	<0.05	0.7055 (0.5477–0.9087)
	Inverse variance weighted	27	<0.05	0.8357 (0.7031–0.9934)
	Simple mode	27	0.6583	0.9053 (0.5854–1.4000)
	Weighted mode	27	<0.05	0.7080 (0.5378–0.9320)
CUB domain–containing protein 1 levels	MR Egger	35	0.0914	1.2981 (0.9673–1.7420)
	Weighted median	35	0.2303	1.1356 (0.9225–1.3979)
	Inverse variance weighted	35	<0.05	1.1922 (1.0140–1.4018)
	Simple mode	35	0.9760	1.0054 (0.7077–1.4284)
	Weighted mode	35	0.1717	1.1446 (0.9470–1.3834)
Interleukin–10 levels	MR Egger	33	<0.05	0.5101 (0.3490–0.7455)
	Weighted median	33	0.1050	0.8034 (0.6165–1.0469)
	Inverse variance weighted	33	<0.05	0.7756 (0.6368–0.9446)
	Simple mode	33	0.9384	1.0249 (0.5522–1.9020)
	Weighted mode	33	0.3876	0.7954 (0.4766–1.3274)
Interleukin–22 receptor subunit alpha–1 levels	MR Egger	21	0.0550	0.5805 (0.3446–0.9778)
	Weighted median	21	<0.05	0.6561 (0.4580–0.9399)
	Inverse variance weighted	21	<0.05	0.7519 (0.5842–0.9676)
	Simple mode	21	0.2277	0.6587 (0.3413–1.2713)
	Weighted mode	21	0.1309	0.6401 (0.3674–1.1151)
Interleukin–5 levels	MR Egger	24	0.5933	1.1115 (0.7584–1.6289)
	Weighted median	24	<0.05	0.7253 (0.5395–0.9752)
	Inverse variance weighted	24	<0.05	0.7891 (0.6494–0.9587)
	Simple mode	24	0.0845	0.5466 (0.2834–1.0540)
	Weighted mode	24	0.4723	1.2139 (0.7217–2.0419)
Leukemia inhibitory factor receptor levels	MR Egger	27	<0.05	0.7115 (0.5330–0.9497)
	Weighted median	27	0.0833	0.8157 (0.6477–1.0272)
	Inverse variance weighted	27	<0.05	0.8224 (0.7049–0.9594)
	Simple mode	27	0.0559	0.6701 (0.4528–0.9918)
	Weighted mode	27	0.0852	0.8050 (0.6347–1.0209)
Osteoprotegerin levels	MR Egger	27	0.1046	0.7210 (0.4927–1.0551)
	Weighted median	27	<0.05	0.6790 (0.5140–0.8969)
	Inverse variance weighted	27	<0.05	0.7975 (0.6631–0.9591)
	Simple mode	27	0.2327	0.7246 (0.4322–1.2148)
	Weighted mode	27	<0.05	0.6715 (0.5005–0.9009)
Stem cell factor levels	MR Egger	42	0.3065	0.8717 (0.6723–1.1303)
	Weighted median	42	0.7797	0.9724 (0.7992–1.1832)
	Inverse variance weighted	42	<0.05	0.8570 (0.7500–0.9793)
	Simple mode	42	0.8201	0.9568 (0.6556–1.3965)
	Weighted mode	42	0.8709	1.0211 (0.7953–1.3109)
Thymic stromal lymphopoietin levels	MR Egger	22	<0.05	1.7988 (1.0761–3.0071)
	Weighted median	22	0.1736	1.2452 (0.9080–1.7077)
	Inverse variance weighted	22	<0.05	1.2610 (1.0095–1.5751)
	Simple mode	22	0.8327	1.0725 (0.5647–2.0372)
	Weighted mode	22	0.8689	1.0488 (0.5997–1.8344)

Fig. 2. Two-step Mendelian randomization (MR) analysis examining the causal relationship between inflammatory cytokines (ICs) and osteomyelitis

ICs. Mendelian randomization analysis showed that all p-values exceeded 0.05, indicating no significant effect of osteomyelitis on circulating IC levels (Supplementary Table 5). Results of horizontal pleiotropy testing, heterogeneity analysis (Supplementary Table 6), and leave-one-out analyses (Supplementary Fig. 3,4) were consistent with these findings.

Metabolic determinants of osteomyelitis risk: Protective and risk-enhancing metabolites

The results presented in Table 1 highlight significant associations between various metabolites and osteomyelitis risk. Using the IVW method, the analysis identified

Table 1. Associations between genetically predicted circulating metabolites and osteomyelitis risk estimated with IVW Mendelian randomization (MR)

Exposure	nsnp	p-value	OR
P-cresol sulfate (PCS)	21	0.012	0.744
X-12007	13	0.007	0.753
X-24307	16	0.018	0.754
Homostachydrine	24	0.007	0.773
X-23648	17	0.038	0.792
Taurodeoxycholate	23	0.012	0.794
Behenoyl sphingomyelin (d18:1/22:0)	18	0.046	0.802
4-hydroxyphenylacetylglutamine	25	0.015	0.809
Pentose acid	22	0.026	0.819
1-palmitoleoyl-2-linolenoyl-GPC (16:1/18:3)	20	0.007	0.823
Citramalate	24	0.019	0.833
Etiocholanolone glucuronide	21	0.012	0.838
N-oleoyltaurine	21	0.023	0.841
Lithocholate sulfate	23	0.037	0.842
Palmitoyl sphingomyelin (d18:1/16:0)	32	0.045	0.847
Creatinine	25	0.038	0.847
Oleoyl-linoleoyl-glycerol (18:1 to 18:2) [2] to linoleoyl-arachidonoyl-glycerol (18:2 to 20:4) [2] ratio	28	0.003	0.850
Hypotaurine	29	0.032	0.853
X-21312	21	0.023	0.854
1-stearoyl-2-linoleoyl-GPI (18:0/18:2)	28	0.029	0.860
Pregnanediol-3-glucuronide	35	0.024	0.861
X-24518	22	0.004	0.862
Adenosine 5′-diphosphate (ADP) to glutamate ratio	30	0.021	0.863
Carnitine	25	0.038	0.873
X-18922	29	0.038	0.873
5-methyluridine (ribothymidine)	26	0.049	0.875
Cytidine	29	0.041	0.885
X-24795	28	0.029	1.094
Gamma-glutamylcitrulline	27	0.047	1.115
2-acetamidophenol sulfate	24	0.050	1.158
Glucuronate to etiocholanolone glucuronide ratio	24	0.041	1.161
Alpha-ketoglutarate to kynurenine ratio	26	0.033	1.172
Tridecenedioate (C13:1-DC)	23	0.041	1.189
N-acetylproline	21	0.047	1.189
Cholesterol to oleoyl-linoleoyl-glycerol (18:1 to 18:2) [2] ratio	35	0.008	1.190
2-naphthol sulfate	23	0.033	1.191
Maltotriose	23	0.012	1.192
Histidine to phosphate ratio	29	0.011	1.204
1-(1-enyl-oleoyl)-GPE (p-18:1)	18	0.022	1.209

Table 1. Associations between genetically predicted circulating metabolites and osteomyelitis risk estimated by inverse-variance weighted (IVW) Mendelian randomization (MR) – cont.

Exposure	nsnp	p-value	OR
Arginine to phosphate ratio	22	0.038	1.210
Glutarate (C5-DC) to salicylate ratio	18	0.043	1.219
1-palmitoyl-2-arachidonoyl-GPI (16:0/20:4)	24	0.015	1.221
1-(1-enyl-stearoyl)-GPE (p-18:0)	24	0.014	1.233
X-18886	16	0.030	1.240
Citrate to oxalate (ethanedioate) ratio	16	0.044	1.242
Alpha-ketobutyrate to 3-methyl-2-oxobutyrate ratio	23	0.014	1.244
Pyruvate	22	0.031	1.252
Tyramine O-sulfate	21	0.017	1.252
X-24978	19	0.004	1.253
Stearidonate (18:4n3)	26	0.033	1.255
Cortisol to 4-cholesten-3-one ratio	20	0.006	1.266
Glycerate	27	0.016	1.269
Carotenoid(cryptoxanthin)	28	0.009	1.334
Taurocholate to oxalate (ethanedioate) ratio	19	0.010	1.361

IVW – inverse-variance weighted (two-sample Mendelian randomization estimator); nsnp – number of single-nucleotide polymorphisms used as instrumental variables; OR – odds ratio per 1-standard-deviation increase in metabolite level; GPC – glycerophosphocholine; GPI – glycerophosphoinositol; GPE – glycerophosphoethanolamine; ADP – adenosine 5′-diphosphate; X-12007, X-23648, X-24307, X-21312, X-24518, X-18922, X-24795, X-18886, X-24978: uncharacterized plasma metabolites designated by their internal genome-wide association study (GWAS) catalogue identifiers.

54 metabolites with either protective or risk-enhancing associations (Supplementary Table 7). Notably, higher levels of PCS, X-12007 (an uncharacterized plasma metabolite identified in the GWAS Catalog), and homostachydrine were associated with a reduced risk of osteomyelitis, with ORs of 0.744 (p = 0.012), 0.753 (p = 0.007), and 0.773 (p = 0.007), respectively. Conversely, higher beta-cryptoxanthin (β-CX) levels were associated with an increased risk of osteomyelitis (OR = 1.334, p = 0.009). Results from pleiotropy and heterogeneity analyses (Supplementary Table 8) were consistent with these findings. An FDR threshold of q < 0.05 was applied separately for each family of tests (cytokines, n = 91; metabolites, n = 1,400; details in Supplementary Table 9).

Mediation roles of ICs on osteomyelitis via metabolites

Previously, we identified 11 ICs with a unidirectional effect on osteomyelitis, as well as 54 metabolites associated with osteomyelitis. We next explored the potential causal relationships between these ICs and metabolites (Table 2). The 95% confidence intervals (95% CIs) for the mediated effect and proportion mediated were estimated using the delta method; proportions >100% or <0% may occur when the indirect and total effects have opposite directions. Figure 3 illustrates the associations between various IC levels and metabolites in relation to osteomyelitis risk, as assessed using different MR methods. The IVW method identified significant associations for specific cytokine–metabolite pairs.

Leukemia inhibitory factor receptor (LIFR) was negatively associated with osteomyelitis, and X-12007 was also

Table 2. Mediation roles of inflammatory cytokines (ICs) on osteomyelitis via metabolites

Pathway	Beta1	Beta2	Beta all	Mediated effect	Mediated proportion
CDCP1 – homostachydrine – osteomyelitis	0.0720123987031509	–0.257209225	0.175824410233942	–0.0185 (–0.067, 0.03)	–10.5% (–38.1%, 17.1%)
LIFR – X-12007 – osteomyelitis	–0.123619426	–0.283756662	–0.195568506	0.0351 (–0.025, 0.0951)	–17.9% (12.8%, –48.6%)
SCF – beta-cryptoxanthin – osteomyelitis	–0.080997106	0.288524677065791	–0.154307188	–0.0234 (–0.086, 0.0392)	15.1% (55.7%, –25.4%)

Beta1 – MR effect (β) of exposure cytokine on mediator metabolite (log-OR or SD units); Beta2 – MR effect (β) of mediator metabolite on outcome (osteomyelitis); Beta all – total MR effect (β) of exposure cytokine on outcome; mediated effect is shown with 95% confidence interval (95% CI) in parentheses; mediated proportion (%) is the percentage of the total cytokine→osteomyelitis effect explained by the metabolite mediator; CDCP1 – CUB domain-containing protein 1; LIFR – leukemia inhibitory factor receptor; SCF – stem cell factor; X-12007, homostachyline, beta-cryptoxanthin – metabolite identifiers/names from the GWAS catalog or Metabolon platform.

exposureTrait	outcomeTrait	nsnp	method	pval	OR (95% CI)
X–12007 levels	Osteomyelitis	13	MR Egger	0.176	0.745 (0.501 to 1.110)
		13	Weighted median	**0.033**	0.738 (0.558 to 0.976)
		13	Inverse variance weighted	**0.007**	0.753 (0.612 to 0.926)
		13	Simple mode	0.121	0.709 (0.473 to 1.062)
		13	Weighted mode	0.144	0.747 (0.518 to 1.077)
Leukemia inhibitory factor receptor levels	X–12007 levels	28	MR Egger	0.408	0.924 (0.767 to 1.112)
		28	Weighted median	0.116	0.906 (0.801 to 1.025)
		28	Inverse variance weighted	**0.015**	0.884 (0.800 to 0.976)
		28	Simple mode	0.113	0.829 (0.663 to 1.038)
		28	Weighted mode	0.070	0.889 (0.786 to 1.004)
Leukemia inhibitory factor receptor levels	Osteomyelitis	27	MR Egger	**0.029**	0.711 (0.533 to 0.950)
		27	Weighted median	0.084	0.816 (0.647 to 1.028)
		27	Inverse variance weighted	**0.013**	0.822 (0.705 to 0.959)
		27	Simple mode	**0.033**	0.670 (0.473 to 0.950)
		27	Weighted mode	0.078	0.805 (0.638 to 1.015)

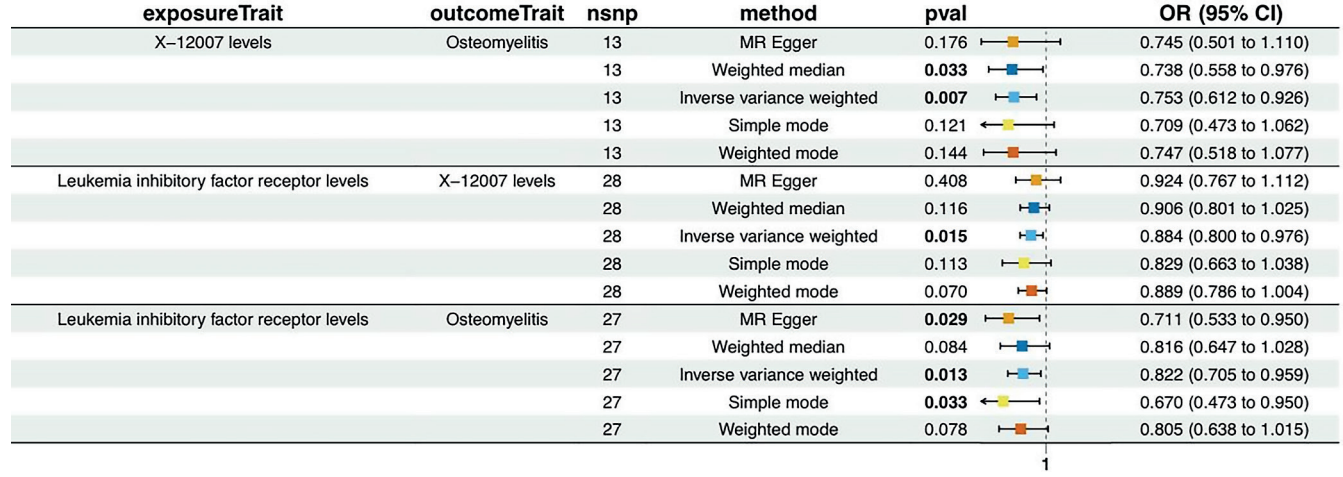

Fig. 3. Forest plots illustrating the associations between inflammatory cytokine (IC) levels, metabolites, and osteomyelitis

associated with a reduced osteomyelitis risk. Based on these findings, one might expect a positive association between LIFR and X-12007. However, the observed causal relationship between LIFR and X-12007 was negative. These findings highlight the complex interplay among cytokines, metabolites, and osteomyelitis risk, with some factors showing protective associations while others may increase susceptibility. Sensitivity analyses further supported these results (Supplementary Table 9 and Supplementary Fig. 5,6).

Validation of osteomyelitis in vitro model

Following establishment of the in vitro osteomyelitis cell model, the expression levels of the inflammatory markers TNF-α and IL-1β in the cell culture supernatant were measured using ELISA (Fig. 4A,B). In the osteomyelitis model, both inflammatory cytokines were significantly elevated compared with the control group. TNF-α concentration was 16.81 ±1.31 pg/mL in the control group and 38.50 ±1.42 pg/mL in the osteomyelitis group (mean ±SD, n = 3 biological replicates; mean difference (MD) = 19.66 pg/mL, 95% CI: 15.75–23.57; unpaired 2-tailed t-test, p = 0.001). Interleukin-1 beta showed a similar trend, with concentrations of 29.82 ±1.36 pg/mL in the control group vs 38.50 ±1.28 pg/mL in the osteomyelitis group (MD = 8.68 pg/mL, 95% CI: 7.32–10.04, p ≤ 0.001).

PCS increased the proliferation and osteogenic differentiation of SPA-treated MC3T3-E1 cells

Furthermore, to explore the impact of PCS on the osteogenic differentiation of SPA-treated MC3T3-E1 cells, ALP activity was evaluated on day 7, and ARS staining was performed on day 21. Cells were seeded at 5×10^4 cells per well in 3 independent cultures (n = 3) and treated with 50 µg mL PCS, the concentration chosen from pilot data. Alkaline phosphatase activity reached 110.5 ±2.1% vs 100.0 ±2.3% (mean ±SD, n = 3 biological replicates); MD = 10.5% (95% CI: 4.8–16.3, p = 0.007). Mineralized nodules (ARS) quantified 131.0 ±1.6% vs 100.0 ±1.8%; MD = 31.0% (95% CI: 26.6–35.4, p < 0.001). The findings demonstrated that MC3T3-E1 cells co-cultured with PCS displayed a notable rise in both ALP activity and calcified nodules compared to the control group (Fig. 5B,C).

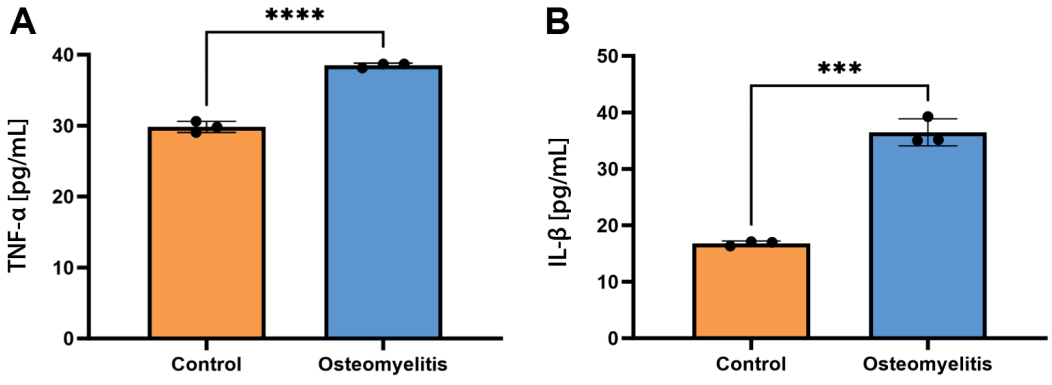

Fig. 4. Enzyme-linked immunosorbent assay (ELISA) quantification of (A) tumor necrosis factor alpha (TNF-α) and (B) interleukin (IL)-1β in cell culture supernatants. Data are presented as mean ± standard deviation (SD) (pg mL) from 3 independent cultures per group. The p-values were calculated using unpaired two-tailed t-tests following Shapiro–Wilk normality testing (p < 0.05, p < 0.01)

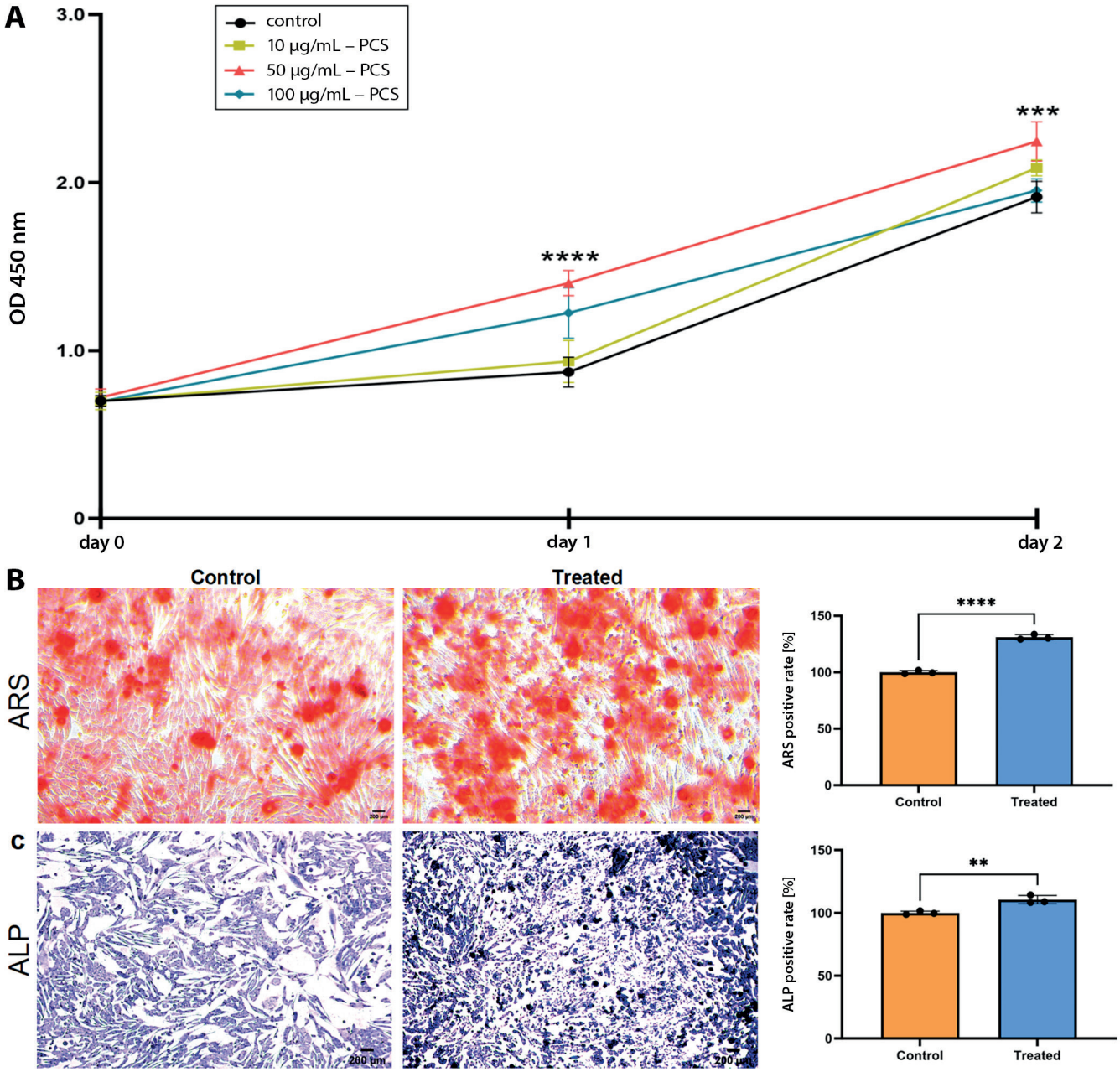

Fig. 5. A. Cell Counting Kit-8 (CCK-8) proliferation assay showing OD_{450} values (mean ± standard deviation (SD)) from 3 biological replicates per group. Dunnett's test vs control yielded an adjusted p = 0.007, with a 95% confidence interval (95% CI) of 4.8–16.3% for the 50 µg/mL PCS group; B. Alkaline phosphatase (ALP) activity (U mg protein) and (C) Alizarin Red S (ARS) absorbance (562 nm) measured after 7 and 21 days, respectively (n = 3). Comparisons were performed using unpaired two-tailed t-tests ($p < 0.05$). Scale bars: 200 µm; magnification: ×200

OD_{450} –optical density at 450 nm; PCS – p-cresol sulfate.

PCS decreased the intracellular ROS production and protein levels of inflammatory cytokines in SPA-treated MC3T3-E1 cells

The reduction effect of PCS on ROS was assessed by observing the fluorescence intensity of the DCFH-DA probe via fluorescence microscopy (Nikon Eclipse 80i) (Fig. 6A). The PCS demonstrated antioxidant properties, significantly reducing intracellular ROS in SPA-treated MC3T3-E1 cells. Intracellular ROS was 1.000 ±0.016 a.u. vs 1.315 ±0.018 a.u. in control (mean ±SD, n = 3 biological replicates); MD = −0.315 a.u. (95% CI: −0.360 to −0.269, $p < 0.001$). Normality and equal variance confirmed ($F = 3.04$, $p = 0.495$). The ELISA results showed that PCS co-culture simultaneously reduced both cytokines: TNF-α decreased to 28.23 ±1.39 pg/mL vs 36.48 ±1.23 pg/mL in SPA-only control (mean ±SD, n = 3 biological replicates; MD = −8.25 pg/mL, 95% CI: −12.21 to −4.29, unpaired two-tailed t-test, $p = 0.004$), while IL-1β fell to 19.60 ±1.21 pg/mL vs 22.36 ±1.18 pg/mL (MD = −2.76 pg/mL, 95% CI: −3.77 to −1.76, $p = 0.002$).

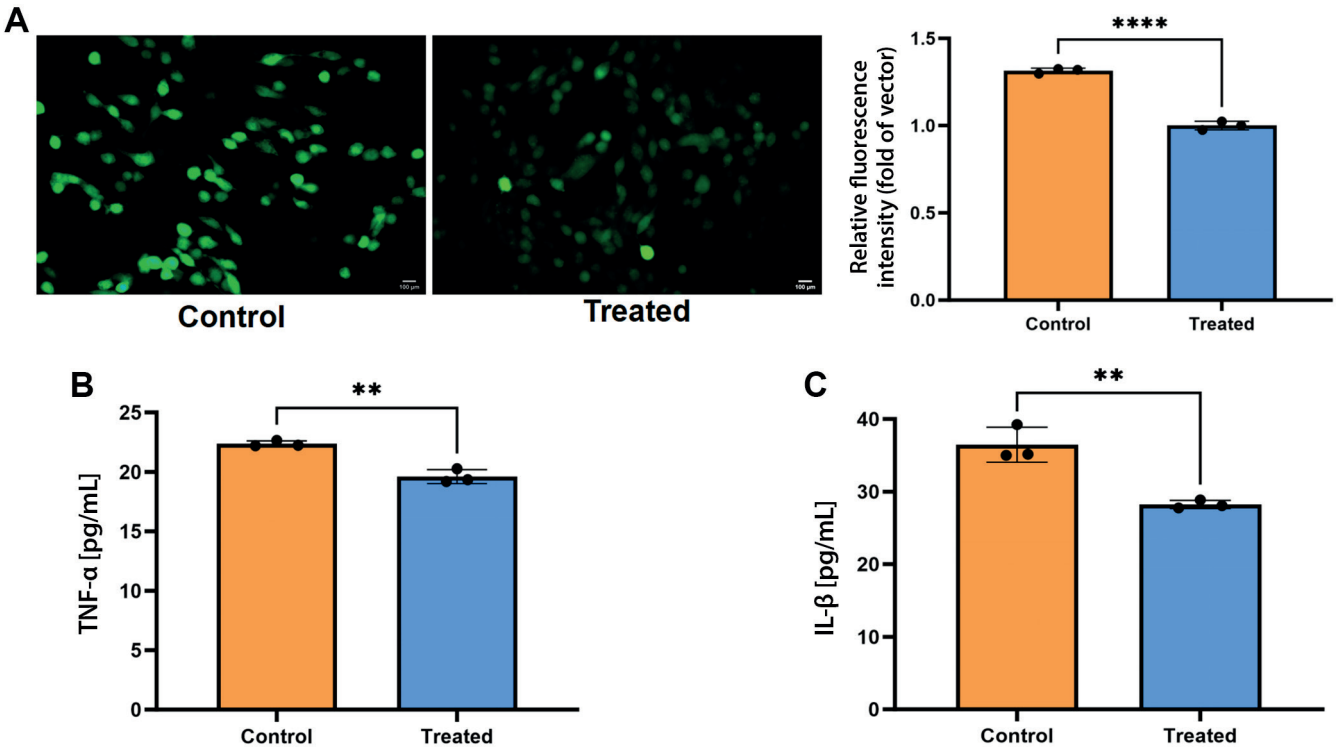

Fig. 6. A. Intracellular reactive oxygen species (ROS) fluorescence (arbitrary units (a.u.), mean ± standard deviation (SD)) from 6 independent cultures per group; B,C. Concentrations of tumor necrosis factor alpha (TNF-α) and interleukin (IL)-1β (pg mL), measured using enzyme-linked immunosorbent assay (ELISA). The p-values were calculated using unpaired two-tailed t-tests with Welch's correction following Shapiro–Wilk normality testing ($p < 0.05$, $p < 0.01$). Scale bars: 100 µm; magnification: ×400

Normality (Shapiro–Wilk) and equal variance (F test) were satisfied for both datasets (TNF-α: F = 18.04, p = 0.105; IL-1β: F = 5.95, p = 0.288). (Fig. 6B,C).

Discussion

Previous research has predominantly focused on the role of ICs in other inflammatory conditions, such as rheumatoid arthritis and inflammatory bowel disease (IBD).[24] In contrast, this study represents the first systematic evaluation of the causal relationship between ICs and osteomyelitis. The results indicate that *CDCP1* and *TSLP* significantly increase the risk of osteomyelitis, whereas *ADA* and *IL-5* appear to exert protective effects. *CDCP1* and *TSLP* may exacerbate the inflammatory response in osteomyelitis by promoting the activation and proliferation of inflammatory cells. *CDCP1* is a transmembrane protein with critical intracellular signaling functions.[25] Its detection on extracellular vesicles in prostate cancer suggests that this cell surface protein may serve as a biomarker for malignancy.[26] As a ligand for CD6 expressed on T cells, *CDCP1* acts as a co-stimulatory receptor within the immunological synapse and has been proposed as a potential molecular target for cancer therapy.[27]

Although *TSLP* was originally identified as an initiator of type 2 inflammatory responses, its biological functions extend far beyond this role, with involvement in viral infections, cancer, chronic inflammatory diseases, and lipid metabolism.[28,29]

In contrast, *ADA* and *IL-5* appear to play protective roles in osteomyelitis by suppressing inflammatory cell activation and attenuating the inflammatory response. *ADA*, an RNA-editing enzyme that acts on RNA, restricts the accumulation of endogenous immunostimulatory double-stranded RNA.[30] The inflammatory effects of *ADA* on tissues may vary depending on the time elapsed since injury onset.[31]

Interleukin-5, a hematopoietic cytokine derived from Th2 cells, plays a critical role in regulating the development and function of basophils and mast cells.[32] Recent studies suggest that *IL-5* may also hold therapeutic potential in sepsis through its effects on non-eosinophilic myeloid cell populations.[33] Ongoing and future clinical trials are expected to further enhance our understanding of its role in the pathogenesis of allergic and eosinophilic inflammatory diseases.[34]

Additionally, this study found that β-CX increases the risk of osteomyelitis, whereas PCS appears to exert protective effects. Residual pleiotropy cannot be entirely excluded, and the restriction of the study population to individuals of European ancestry limits the generalizability of these findings to other populations. Compared with the MR analysis by Yang et al.,[35] our study expands the metabolic landscape to include more than 1,400 metabolites and provides additional functional validation.

As a carotenoid, β-CX is generally recognized for its anti-inflammatory properties, particularly in conditions such as oral ulcers, where it suppresses inflammation-related signaling pathways and cytokine production.[36] However, in the context of osteomyelitis, β-CX appears to exhibit pro-inflammatory effects. This apparent discrepancy may be attributable to differences in cellular and tissue microenvironments, as well as the specific signaling pathways involved. For example, β-CX has been shown to enhance nuclear *Nrf2* expression and preserve mitochondrial function, thereby mitigating oxidative stress and cellular senescence in renal cells; however, within the bone marrow microenvironment, it may activate distinct molecular pathways that promote inflammation.[37]

PCS is the principal metabolite of p-cresol, a uremic toxin traditionally associated with toxicity and increased mortality in patients with chronic kidney disease (CKD).[38] Previous studies have demonstrated that PCS induces immune activation and cellular inflammatory responses in various cell types, including cultured proximal renal tubular cells, endothelial cells, and macrophages.[39]

However, the relationship between PCS and its precursor, tyrosine, appears to be complex. Notably, PCS levels are not correlated with tyrosine concentrations, despite evidence that tyrosine can suppress the production of pro-inflammatory ICs.[40] Our MR-prioritized identification of PCS is consistent with recent murine metabolomics studies that identified sphingosine and tricarboxylic acid (TCA) cycle intermediates as early indicators of osteomyelitis.[41] These findings suggest that the inflammatory effects of PCS may be mediated not by its precursor itself, but rather by downstream metabolic mechanisms. Given this, p-cresol and its metabolites represent promising targets for future mechanistic investigation.[42]

In the context of osteomyelitis, the protective effects of PCS may be attributable to its potential modulation of oxidative stress pathways. Oxidative stress is a critical contributor to the pathogenesis of osteomyelitis, as it can exacerbate inflammation and tissue damage.[43] By modulating the cellular oxidative stress response, PCS may help attenuate inflammatory cell activation, thereby reducing disease severity. This hypothesis is supported by studies demonstrating that PCS can influence the expression of antioxidant enzymes, such as superoxide dismutase and catalase, which play key roles in neutralizing ROS.[39] Our findings identify PCS as a potential candidate for future mechanistic and interventional studies.

Limitations of the study

This study has several limitations. First, the analysis was restricted to GWAS data derived from individuals of European ancestry, which may limit the generalizability of the findings to other populations. Second, residual pleiotropy and weak instrumental variables could not be entirely excluded. Third, the mediation analyses were based on genetic proxies rather than direct measurements of circulating biomarker levels. Fourth, the in vitro validation was limited to a single murine cell line and short-term experimental readouts. Finally, potential confounding factors, including age, sex, and environmental exposures, could not be accounted for in the summary-level data.

Conclusions

Despite leveraging comprehensive GWAS datasets, our investigation was predominantly limited to populations of European ancestry and lacked critical demographic information, such as age and sex distribution, thereby precluding more refined stratified analyses that might have identified population-specific associations.[44] In summary, our MR analysis identified PCS as a potentially protective metabolite against osteomyelitis (OR = 0.74, p = 0.012), a finding further supported by in vitro evidence demonstrating reduced ROS production and lower levels of pro-inflammatory cytokines. Future studies involving diverse ancestral populations and well-controlled experimental models are warranted to validate and further elucidate these findings.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.17622218. The package contains the following files:

Supplementary Table 1. Instrumental variables for inflammatory cytokines.

Supplementary Table 2. SNPs used as instruments.

Supplementary Table 3. Instrumental variables for osteomyelitis.

Supplementary Table 4. Heterogeneity and pleiotropy tests for total effects (beta.all).

Supplementary Table 5. All odds ratios for nonsignificant cytokines (all_OR_NO10).

Supplementary Table 6. Heterogeneity and pleiotropy tests for reverse MR analysis.

Supplementary Table 7. IVW filtered results (IVW.filter).

Supplementary Table 8. Heterogeneity and pleiotropy tests for beta2 (metabolite→outcome).

Supplementary Table 9. Heterogeneity and pleiotropy tests for beta1 (cytokine→metabolite).

Data Availability Statement

The datasets supporting the findings of this study are available in 2 ways:

1. Publicly available GWAS summary statistics used as exposure and outcome were accessed from FinnGen R10 (https://storage.googleapis.com/finngen-public-data-r10/summary_stats/finngen_R10_M13_OSTEOMYELITIS.gz)

and GWAS Catalog (GCST90274758–GCST90274848 for inflammatory proteins; GCST90199621–GCST90201020 for metabolites).

2. All raw and processed data are openly available in Zenodo at https://doi.org/10.5281/zenodo.17622259.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Xingyu Chen https://orcid.org/0009-0006-3673-9124
Ruiqing Mo https://orcid.org/0009-0008-9088-0203
Sijie Yang https://orcid.org/0009-0007-9535-124X
Peilin Zhou https://orcid.org/0009-0008-4980-0582
Hua Qikai https://orcid.org/0009-0006-3673-9124

References

1. Wright JA, Nair SP. Interaction of Staphylococci with bone. *Int J Med Microbiol*. 2010;300(2–3):193–204. doi:10.1016/j.ijmm.2009.10.003
2. Mörmann M, Thederan M, Nackchbandi I, Giese T, Wagner C, Hänsch GM. Lipopolysaccharides (LPS) induce the differentiation of human monocytes to osteoclasts in a tumour necrosis factor (TNF) α-dependent manner: A link between infection and pathological bone resorption. *Mol Immunol*. 2008;45(12):3330–3337. doi:10.1016/j.molimm.2008.04.022
3. Wang X, Zhang M, Zhu T, Wei Q, Liu G, Ding J. Flourishing antibacterial strategies for osteomyelitis therapy. *Adv Sci (Weinh)*. 2023;10(11):2206154. doi:10.1002/advs.202206154
4. Masters EA, Ricciardi BF, Bentley KLDM, Moriarty TF, Schwarz EM, Muthukrishnan G. Skeletal infections: Microbial pathogenesis, immunity and clinical management. *Nat Rev Microbiol*. 2022;20(7):385–400. doi:10.1038/s41579-022-00686-0
5. Kavanagh N, Ryan EJ, Widaa A, et al. Staphylococcal osteomyelitis: Disease progression, treatment challenges, and future directions. *Clin Microbiol Rev*. 2018;31(2):e00084-17. doi:10.1128/CMR.00084-17
6. Mödinger Y, Löffler B, Huber-Lang M, Ignatius A. Complement involvement in bone homeostasis and bone disorders. *Semin Immunol*. 2018;37:53–65. doi:10.1016/j.smim.2018.01.001
7. Funao H, Ishii K, Nagai S, et al. Establishment of a real-time, quantitative, and reproducible mouse model of *Staphylococcus* osteomyelitis using bioluminescence imaging. *Infect Immun*. 2012;80(2):733–741. doi:10.1128/IAI.06166-11
8. Hofmann SR, Kapplusch F, Girschick HJ, et al. Chronic recurrent multifocal osteomyelitis (CRMO): Presentation, pathogenesis, and treatment. *Curr Osteoporos Rep*. 2017;15(6):542–554. doi:10.1007/s11914-017-0405-9
9. Yu T, Bai R, Wang Z, et al. Colon-targeted engineered postbiotics nanoparticles alleviate osteoporosis through the gut–bone axis. *Nat Commun*. 2024;15(1):10893. doi:10.1038/s41467-024-55263-1
10. Guo J, Wang F, Hu Y, et al. Exosome-based bone-targeting drug delivery alleviates impaired osteoblastic bone formation and bone loss in inflammatory bowel diseases. *Cell Rep Med*. 2023;4(1):100881. doi:10.1016/j.xcrm.2022.100881
11. Safiri S, Kolahi AA, Cross M, et al. Global, regional, and national burden of other musculoskeletal disorders 1990–2017: Results from the Global Burden of Disease Study 2017. *Rheumatology (Oxford)*. 2021;60(2):855–865. doi:10.1093/rheumatology/keaa315
12. Yang Z, Lin B, Ren H, et al. Risk factors for osteomyelitis: A systematic review and meta-analysis. *Int J Surg*. 2025;111(8):5606–5622. doi:10.1097/JS9.0000000000002811
13. Davies NM, Holmes MV, Davey Smith G. Reading Mendelian randomisation studies: A guide, glossary, and checklist for clinicians. *BMJ*. 2018;362:k601. doi:10.1136/bmj.k601
14. Gehrke AKE, Mendoza-Bertelli A, Ledo C, et al. Neutralization of *Staphylococcus aureus* protein A prevents exacerbated osteoclast activity and bone loss during osteomyelitis. *Antimicrob Agents Chemother*. 2023;67(1):e01140-22. doi:10.1128/aac.01140-22
15. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40(4):304–314. doi:10.1002/gepi.21965
16. Hartwig FP, Davey Smith G, Bowden J. Robust inference in summary data Mendelian randomization via the zero modal pleiotropy assumption. *Int J Epidemiol*. 2017;46(6):1985–1998. doi:10.1093/ije/dyx102
17. Sun BB, Chiou J, Traylor M, et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023;622(7982):329–338. doi:10.1038/s41586-023-06592-6
18. Diemer EW, Labrecque JA, Neumann A, Tiemeier H, Swanson SA. Mendelian randomisation approaches to the study of prenatal exposures: A systematic review. *Paediatr Perinat Epidemiol*. 2021;35(1):130–142. doi:10.1111/ppe.12691
19. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife*. 2018;7:e34408. doi:10.7554/eLife.34408
20. Wang J, Li J, Ji Y. Mendelian randomization as a cornerstone of causal inference for gut microbiota and related diseases from the perspective of bibliometrics. *Medicine (Baltimore)*. 2024;103(26):e38654. doi:10.1097/MD.0000000000038654
21. Liang W, Li Y, Ji Y, et al. Exosomes derived from bone marrow mesenchymal stem cells induce the proliferation and osteogenic differentiation and regulate the inflammatory state in osteomyelitis in vitro model. *Naunyn Schmiedebergs Arch Pharmacol*. 2025;398(2):1695–1705. doi:10.1007/s00210-024-03357-4
22. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018;50(5):693–698. doi:10.1038/s41588-018-0099-7
23. Sanderson E, Davey Smith G, Windmeijer F, Bowden J. An examination of multivariable Mendelian randomization in the single-sample and two-sample summary data settings. *Int J Epidemiol*. 2019;48(3):713–727. doi:10.1093/ije/dyy262
24. Neurath MF. Strategies for targeting cytokines in inflammatory bowel disease. *Nat Rev Immunol*. 2024;24(8):559–576. doi:10.1038/s41577-024-01008-6
25. Noren Hooten N, Torres S, Mode NA, et al. Association of extracellular vesicle inflammatory proteins and mortality. *Sci Rep*. 2022;12(1):14049. doi:10.1038/s41598-022-17944-z
26. Minciacchi VR, Zijlstra A, Rubin MA, Di Vizio D. Extracellular vesicles for liquid biopsy in prostate cancer: Where are we and where are we headed? *Prostate Cancer Prostatic Dis*. 2017;20(3):251–258. doi:10.1038/pcan.2017.7
27. Kollmorgen G, Niederfellner G, Lifke A, et al. Antibody mediated CDCP1 degradation as mode of action for cancer targeted therapy. *Mol Oncol*. 2013;7(6):1142–1151. doi:10.1016/j.molonc.2013.08.009
28. Ebina-Shibuya R, Leonard WJ. Role of thymic stromal lymphopoietin in allergy and beyond. *Nat Rev Immunol*. 2023;23(1):24–37. doi:10.1038/s41577-022-00735-y
29. Verstraete K, Peelman F, Braun H, et al. Structure and antagonism of the receptor complex mediated by human TSLP in allergy and asthma. *Nat Commun*. 2017;8(1):14937. doi:10.1038/ncomms14937
30. De Reuver R, Verdonck S, Dierick E, et al. ADAR1 prevents autoinflammation by suppressing spontaneous ZBP1 activation. *Nature*. 2022;607(7920):784–789. doi:10.1038/s41586-022-04974-w
31. Bagheri S, Saboury AA, Haertlé T. Adenosine deaminase inhibition. *Int J Biol Macromol*. 2019;141:1246–1257. doi:10.1016/j.ijbiomac.2019.09.078
32. Molfino NA, Gossage D, Kolbeck R, Parker JM, Geba GP. Molecular and clinical rationale for therapeutic targeting of interleukin-5 and its receptor. *Clin Exp Allergy*. 2012;42(5):712–737. doi:10.1111/j.1365-2222.2011.03854.x
33. Linch SN, Danielson ET, Kelly AM, Tamakawa RA, Lee JJ, Gold JA. Interleukin 5 is protective during sepsis in an eosinophil-independent manner. *Am J Respir Crit Care Med*. 2012;186(3):246–254. doi:10.1164/rccm.201201-0134OC

34. Dougan M, Dranoff G, Dougan SK. GM-CSF, IL-3, and IL-5 family of cytokines: Regulators of inflammation. *Immunity*. 2019;50(4):796–811. doi:10.1016/j.immuni.2019.03.022
35. Chai W, Yuan H, Liu J, Yang Y. Systematic proteome-wide Mendelian randomization using the human plasma proteome to identify therapeutic targets for osteomyelitis [published online as ahead of print on July 14, 2025]. *Naunyn Schmiedebergs Arch Pharmacol*. 2025. doi:10.1007/s00210-025-04429-9
36. Yamanobe H, Yamamoto K, Kishimoto S, et al. Anti-inflammatory effects of β-cryptoxanthin on 5-fluorouracil-induced cytokine expression in human oral mucosal keratinocytes. *Molecules*. 2023;28(7):2935. doi:10.3390/molecules28072935
37. Ke J, Zang H, Liu Y, et al. β-cryptoxanthin suppresses oxidative stress via activation of the Nrf2/HO-1 signaling pathway in diabetic kidney disease. *Front Pharmacol*. 2024;15:1480629. doi:10.3389/fphar.2024.1480629
38. Rong Y, Kiang TKL. Characterization of human sulfotransferases catalyzing the formation of p-cresol sulfate and identification of mefenamic acid as a potent metabolism inhibitor and potential therapeutic agent for detoxification. *Toxicol Appl Pharmacol*. 2021;425:115553. doi:10.1016/j.taap.2021.115553
39. Jing YJ, Ni JW, Ding FH, et al. p-Cresyl sulfate is associated with carotid arteriosclerosis in hemodialysis patients and promotes atherogenesis in apoE$^{-/-}$ mice. *Kidney Int*. 2016;89(2):439–449. doi:10.1038/ki.2015.287
40. Shipelin VA, Trusov NV, Apryatin SA, et al. Effects of tyrosine and tryptophan in rats with diet-induced obesity. *Int J Mol Sci*. 2021;22(5):2429. doi:10.3390/ijms22052429
41. Isogai N, Shiono Y, Kuramoto T, et al. Potential osteomyelitis biomarkers identified by plasma metabolome analysis in mice. *Sci Rep*. 2020;10(1):839. doi:10.1038/s41598-020-57619-1
42. Paul KC, Zhang K, Walker DI, et al. Untargeted serum metabolomics reveals novel metabolite associations and disruptions in amino acid and lipid metabolism in Parkinson's disease. *Mol Neurodegener*. 2023;18(1):100. doi:10.1186/s13024-023-00694-5
43. Fan M, Ren Y, Zhu Y, et al. Borosilicate bioactive glass synergizing low-dose antibiotic loaded implants to combat bacteria through ATP disruption and oxidative stress to sequentially achieve osseointegration. *Bioact Mater*. 2025;44:184–204. doi:10.1016/j.bioactmat.2024.10.009
44. Chen W, Wu Y, Zheng Z, et al. Improved analyses of GWAS summary statistics by reducing data heterogeneity and errors. *Nat Commun*. 2021;12(1):7117. doi:10.1038/s41467-021-27438-7

Causal links between gut microbiota, metabolites, and diffuse large B-cell lymphoma: Evidence from a Mendelian randomization study

Ganyu Feng[1,A–D], Guangcai Zhong[1,A,B], Cong Wang[1,B], Ran Kong[1,B], *Jianhong Wang[1,A,E], *Xiangxiang Zhou[1,2,A,E,F]

[1] Department of Hematology, Shandong Provincial Hospital Affiliated to Shandong First Medical University, Jinan, China
[2] Department of Hematology, Shandong Provincial Hospital, Shandong University, Jinan, 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(5):877–888

Address for correspondence
Xiangxiang Zhou
E-mail: xiangxiangzhou@sdu.edu.cn

Funding sources
None declared

Conflict of interest
None declared

*Jianhong Wang and Xiangxiang Zhou contributed equally to this work.

Received on March 9, 2025
Reviewed on May 19, 2025
Accepted on July 28, 2025

Published online on March 25, 2026

Cite as
Feng G, Zhong G, Wang C, Kong R, Wang J, Zhou X. Causal links between gut microbiota, metabolites, and diffuse large B-cell lymphoma: Evidence from a Mendelian randomization study. *Adv Clin Exp Med*. 2026;35(5):877–888. doi:10.17219/acem/208704

DOI
10.17219/acem/208704

Abstract

Background. The precise causal relationship between alterations in the gut microbiota, microbiota-derived metabolites, and the development of diffuse large B-cell lymphoma (DLBCL) remains unclear.

Objectives. To investigate the potential causal relationships between gut microbiota, microbiota-derived metabolites, and DLBCL.

Materials and methods. Genetic data on gut microbiota were obtained from the MiBioGen consortium, while data on microbiota-derived metabolites were sourced from the TwinsUK and KORA studies. Summary statistics for DLBCL were retrieved from FinnGen. Mendelian randomization (MR) analysis was performed, with inverse-variance weighting (IVW) used as the primary analytical method. Sensitivity analyses included Cochran's Q test, the MR-Egger intercept test, and MR-PRESSO. Reverse MR analysis was conducted to assess potential bidirectional causal relationships between gut microbiota and DLBCL. Bayesian weighted MR (BWMR) was applied for additional validation to enhance the robustness of the findings.

Results. Among 196 gut microbial taxa analyzed, *Bilophila* (odds ratio (OR) = 1.777, 95% confidence interval (95% CI): 1.053–3.000, p = 0.031) was associated with an increased risk of DLBCL. In contrast, *Alistipes* (OR = 0.521, 95% CI: 0.311–0.873, p = 0.013) and *Ruminococcaceae* UCG011 (OR = 0.749, 95% CI: 0.574–0.978, p = 0.034) were associated with a reduced risk. Reverse MR analysis demonstrated a positive association between DLBCL risk and the abundance of *Anaerofilum* (OR = 1.087, 95% CI: 1.008–1.173, p = 0.031). Negative associations were observed between DLBCL risk and the abundance of Deltaproteobacteria (OR = 0.959, 95% CI: 0.922–0.997, p = 0.037), *Desulfovibrionales* (OR = 0.959, 95% CI: 0.922–0.998, p = 0.041), Oxalobacteraceae (OR = 0.914, 95% CI: 0.843–0.992, p = 0.031), and *Oxalobacter* (OR = 0.909, 95% CI: 0.837–0.988, p = 0.024). Analysis of microbiota-derived metabolites identified a causal association between indolepropionate (OR = 0.296, 95% CI: 0.131–0.669, p = 0.003) and reduced DLBCL risk, whereas 7-alpha-hydroxy-3-oxo-4-cholestenoate (7-HOCA) (OR = 9.561, 95% CI: 1.426–64.088, p = 0.020) was associated with an increased risk. No evidence of directional pleiotropy or heterogeneity was detected.

Conclusions. This MR study provides evidence that specific gut microbial taxa and microbiota-derived metabolites may causally influence the risk of DLBCL.

Key words: metabolomics, gut microbiota, diffuse large B-cell lymphoma, Mendelian randomization, Bayesian statistics

Highlights

- Mendelian randomization identifies causal associations between gut microbiota, microbial metabolites, and diffuse large B-cell lymphoma (DLBCL) risk.
- *Alistipes* and *Ruminococcaceae* UCG011 are inversely associated with DLBCL, whereas *Bilophila* increases disease risk.
- The microbial metabolite indolepropionate shows protective effects against DLBCL, while 7-HOCA is positively associated with lymphoma development.
- Evidence supports bidirectional interactions between DLBCL and specific gut bacterial taxa.

Background

Non-Hodgkin lymphoma (NHL) comprises a heterogeneous group of malignant lymphoproliferative disorders arising from lymphocytes.[1] In recent decades, the incidence of NHL has increased significantly, and it now ranks as the 11th most common malignancy worldwide, contributing substantially to the global cancer burden.[1,2] Diffuse large B-cell lymphoma (DLBCL) is the most prevalent subtype of NHL, accounting for approx. 30–58% of cases, with notable geographic variation in its distribution.[3,4]

Although DLBCL is considered clinically aggressive, standard treatment regimens generally yield favorable outcomes, with 60–70% of patients achieving a 5-year overall survival (OS) rate. Nevertheless, approx. 30–40% of patients eventually experience relapse or treatment failure, partly due to an incomplete understanding of the disease's underlying pathogenic mechanisms.[3]

Although the etiology of DLBCL remains incompletely elucidated, accumulating evidence suggests its development involves complex interactions between genetic susceptibility, immune dysregulation, and environmental exposures.[5] In this context, the proposed microbiota–gut–lymphoma axis highlights the potential mechanistic link between microbial alterations and lymphomagenesis.[6,7] The human gastrointestinal tract contains approx. 100 trillion microorganisms,[8,9] which contribute to host physiology by enhancing nutrient absorption,[10] providing pathogen defense,[11] maintaining intestinal barrier function, modulating epithelial differentiation,[12] and regulating immune responses.[13] However, disruption of these microbial communities, referred to as dysbiosis, may impair normal physiological functions.[9] Characterized by a reduction in beneficial microbes and/or an increase in pathogenic species, dysbiosis has been linked to chronic inflammation, immune dysregulation, and tumor development.[7,14] Emerging evidence indicates that the gut microbiota contributes not only to the development of gastrointestinal malignancies but also to oncogenesis at extraintestinal sites.[14–19] Recent studies have identified distinct gut microbial profiles in newly diagnosed DLBCL patients compared with healthy controls,[18,20–22] with specific alterations in microbial composition correlating with clinical outcomes.[22,23] Furthermore, accumulating evidence indicates that microbial metabolites may influence DLBCL progression by modulating immunometabolic pathways.[24]

Although current research has demonstrated correlations between gut microbial profiles, microbial metabolites, and DLBCL, causal relationships remain unproven, as most existing studies have employed case–control designs and therefore cannot establish temporal relationships between microbial alterations and lymphomagenesis.[22] In addition, observational studies examining gut microbiota–DLBCL associations are subject to inherent limitations due to multiple confounding factors, including age, environmental exposures, diet, lifestyle, comorbidities, and therapeutic interventions, which may substantially bias association estimates.[25] While randomized controlled trials represent the gold standard for causal inference, their implementation is often limited by ethical and practical constraints. Therefore, Mendelian randomization (MR) and other causal inference methodologies may offer alternative approaches to elucidate potential microbiome–lymphoma relationships. Mendelian randomization is an epidemiological method that strengthens causal inference by using genetic variants as instrumental variables (IVs), leveraging the random allocation of alleles at conception to minimize confounding from postnatal environmental factors.[26] This approach has gained increasing utility in investigating potential causal links between gut microbiota composition and various cancers.[27,28]

This study employed a two-sample MR framework to systematically investigate potential causal relationships between gut microbiota composition and DLBCL risk. We aimed to elucidate the role of specific gut microbial taxa in DLBCL pathogenesis while simultaneously assessing the bidirectional nature of microbiota–lymphoma associations. Furthermore, we evaluated potential causal relationships between gut microbiota–derived metabolites and DLBCL development. By applying robust MR methodologies, our findings may provide novel insights into gut microbiome–DLBCL interactions and potentially inform the development of microbiota-targeted strategies for DLBCL prevention and clinical management.

Objectives

This study employed MR to systematically investigate causal relationships between gut microbiota composition and DLBCL risk, with a particular focus on identifying potentially pathogenic microbial taxa and their contributions to DLBCL lymphomagenesis. Our findings may provide insights into the development of microbiome-modulating interventions for DLBCL prevention and treatment.

Materials and methods

Study design

This MR study was conducted in accordance with the Epidemiology-MR (STROBE-MR) guidelines. We applied both conventional two-sample MR and Bayesian weighted MR (BWMR) to assess causal relationships between gut microbiota composition, microbial metabolites, and DLBCL risk. A schematic overview of the study design is presented in Fig. 1. Instrumental variables were selected as single nucleotide polymorphisms (SNPs) significantly associated with the exposures from genome-wide association study (GWAS) datasets. Mendelian randomization analyses were performed using the TwoSampleMR package in the R statistical environment (R Foundation for Statistical Computing, Vienna, Austria). Comprehensive sensitivity analyses, including heterogeneity tests, pleiotropy assessments, and leave-one-out analyses, were conducted to evaluate the robustness of the findings. In addition, BWMR provided further validation of the causal estimates through its robust Bayesian framework. Reverse MR analyses were also performed to evaluate potential reverse causality between gut microbiota and DLBCL. Our analytical framework satisfied the 3 core assumptions of MR:

1. The SNPs used as IVs must be robustly associated with the exposure;
2. The IVs must be independent of confounders influencing both the exposure and the outcome;
3. The IVs must affect the outcome exclusively through the exposure (i.e., absence of horizontal pleiotropy).

Data sources

The gut microbiota dataset was obtained from the MiBioGen consortium GWAS (https://mibiogen.gcc.rug.nl), comprising 18,340 individuals of predominantly European ancestry across 24 cohorts.[29] This resource provided 16S rRNA gene sequencing profiles paired with genome-wide genotyping data, identifying 211 bacterial taxa across 5 taxonomic levels: 9 phyla, 16 classes, 20 orders, 35 families, and 131 genera. To enhance analytical robustness, we excluded 3 taxonomically undefined families and 12 unclassified genera that could introduce annotation

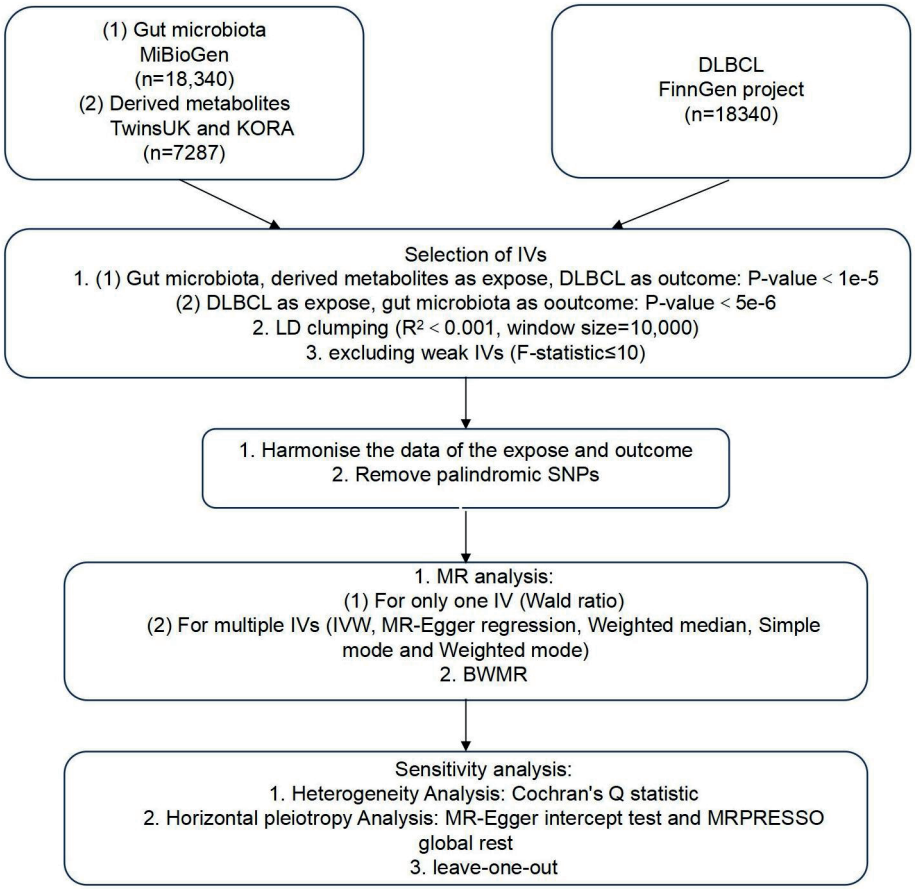

Fig. 1. Flowchart of the Mendelian randomization (MR) analysis investigating the associations between gut microbiota, microbiota-derived metabolites, and diffuse large B-cell lymphoma (DLBCL)

inconsistencies or noise into SNP-exposure associations. Metabolite GWAS data were obtained from 2 European population-based studies (TwinsUK and KORA), comprising 7,824 participants with quantitative profiles of 486 circulating metabolites. From this dataset, we systematically identified and extracted 77 gut microbiota-derived metabolites through annotation based on the Human Metabolome Database (https://hmdb.ca).

The DLBCL summary statistics were obtained from the latest release (R10; published December 18, 2023) of the FinnGen consortium, specifically from the "Diffuse large B-cell lymphoma" dataset, comprising 1,050 cases and 314,193 controls. As this study exclusively utilized de-identified, publicly available data from established consortia, no additional institutional ethics approval was required. When treating gut microbiota and their derived metabolites as exposures and DLBCL as the outcome, the selection of IVs was based on the following criteria: 1) Based on prior literature, we set the significance threshold at $p < 1.0 \times 10^{-5}$ to maximize the number of genetic instruments for exposures[30]; 2) To ensure independence between IVs, LD pruning was conducted using a threshold of $r^2 < 0.001$ and a clumping window of 10,000 kb[30]; 3) Exposure and outcome datasets were harmonized, with palindromic SNPs removed. Single nucleotide polymorphisms with F-statistics ≤10 were excluded to ensure instrument strength, as the F-statistic reflects the strength of the association between IVs and the exposure.[31] For the reverse MR, in which DLBCL was treated as the exposure and gut microbiota as the outcome, the significance threshold was set at $p < 5 \times 10^{-6}$. All other IV selection criteria were consistent with those applied in the forward MR analysis.

Statistical analyses

A range of MR methodologies was applied to investigate the potential causal association between gut microbiota and DLBCL. For bacterial genera with only a single available IV, the Wald ratio method was used to estimate causal effects.[32] When multiple IVs were available, 5 complementary MR methods were applied: inverse-variance weighted (IVW), MR-Egger regression, weighted median, simple mode, and weighted mode. Under the assumption of no horizontal pleiotropy, the IVW method provides the most precise estimates and was therefore selected as the primary analytical approach in this study.[33] MR-Egger regression, which assumes that instrument strength is independent of direct effects (InSIDE assumption), was used as a sensitivity analysis to assess horizontal pleiotropy.[34] The weighted median method can provide consistent causal estimates even when up to 50% of the genetic instruments are invalid.[35] A key excerpt of the R script used for the data analysis is provided in the shared raw data. The F-statistic for each SNP was calculated using the formula $F = \beta^2/SE^2$, and SNPs with F-statistics ≤10 were excluded to minimize weak instrument bias. To assess the robustness of MR estimates, several sensitivity analyses were subsequently performed. Cochran's Q test was used to detect heterogeneity among IVs, with $p < 0.05$ indicating statistical significance.[36] Horizontal pleiotropy was evaluated using both the MR-Egger intercept test (estimating the average pleiotropic effect) and the MR-PRESSO global test (detecting outlier SNPs).[34,37] In addition, a leave-one-out analysis was conducted to determine whether any single IV had a disproportionate influence on causal effect estimates.[38] The BWMR was employed to account for uncertainties arising from polygenicity and to address violations of IV assumptions by controlling for outliers.[39]

Results

MR results

The IVW method was used as the primary analytical approach, with BWMR providing complementary validation. The IVW analysis identified 4 gut microbial taxa showing statistically significant associations with DLBCL risk ($p < 0.05$). Specifically, *Bilophila* (odds ratio (OR) = 1.777, 95% confidence interval (95% CI): 1.053–3.000, p = 0.031) and Desulfovibrionaceae (OR = 1.577, 95% CI: 1.003–2.487, p = 0.049) were positively associated with an increased risk of DLBCL. In contrast, *Alistipes* (OR = 0.521, 95% CI: 0.311–0.873, p = 0.013) and *Ruminococcaceae* UCG011 (OR = 0.749, 95% CI: 0.574–0.978, p = 0.034) were inversely associated with DLBCL risk (Table 1). Scatter plots illustrating the nominally significant associations ($p < 0.05$) identified through IVW analysis are presented in Fig. 2A–C.

The BWMR validation confirmed 3 of these associations and refined their effect estimates. *Alistipes* (OR = 0.573, 95% CI: 0.341–0.964; p = 0.036) and *Ruminococcaceae* UCG011 (OR = 0.747, 95% CI: 0.565–0.987; p = 0.040) maintained their inverse associations with DLBCL risk, whereas *Bilophila* demonstrated a stronger positive association (OR = 2.005, 95% CI: 1.238–3.248; p = 0.005). In contrast, the association for Desulfovibrionaceae was no longer statistically significant in the BWMR analysis (OR = 1.342, 95% CI: 0.846–2.130; p = 0.212), and was therefore not considered in subsequent analyses (Table 1).

Moreover, our MR analysis of gut microbiota–derived metabolites identified 3 significant associations with DLBCL risk. Two metabolites, 3-(4-hydroxyphenyl)lactate (OR = 0.174, 95% CI: 0.030–0.784; p = 0.023) and indolepropionate (OR = 0.296, 95% CI: 0.131–0.669; p = 0.003), were inversely associated with DLBCL risk, suggesting potential protective effects. In contrast, 7α-hydroxy-3-oxo-4-cholestenoate (7-HOCA) (OR = 9.561, 95% CI: 1.426–64.088; p = 0.020) exhibited a strong positive association with DLBCL risk, indicating a potential risk-promoting role (Table 2). BWMR validation confirmed 2 of these associations. Indolepropionate (OR = 0.256, 95% CI:

Table 1. The causal relationship estimation between gut microbiota and diffuse large B-cell lymphoma (DLBCL)

Gut microbiota (exposure)	Number of SNPs	β	p-value	OR (95% CI)	p-value for heterogeneity test	p-value for intercept	p-value for MR-PRESSO
Alistipes							
IVW	13	–0.652	0.013	0.521 (0.311–0.873)	0.889	0.257	0.797
MR Egger	13	0.834	0.525	2.302 (0.190–27.832)	0.928	–	–
Weighted median	13	–0.588	0.088	0.556 (0.283–1.092)	–	–	–
Simple mode	13	–0.498	0.395	0.608 (0.201–1.840)	–	–	–
Weighted mode	13	–0.515	0.369	0.597 (0.202–1.767)	–	–	–
BWMR	13	–0.557	0.036	0.573 (0.341–0.964)	–	–	–
Biophila							
IVW	13	0.575	0.031	1.777 (1.053–3.0)	0.121	0.659	0.089
MR Egger	13	1.181	0.405	3.256 (0.225–47.225)	0.095	–	–
Weighted median	13	0.827	0.012	2.286 (1.192–4.383)	–	–	–
Simple mode	13	1.194	0.092	3.200 (0.870–12.517)	–	–	–
Weighted mode	13	1.183	0.091	3.266 (1.019–10.468)	–	–	–
BWMR	13	0.696	0.005	2.005 (1.238–3.248)	–	–	–
Desulfovibrionaceae							
IVW	10	0.457	0.049	1.579 (1.003–2.487)	0.701	0.271	0.485
MR Egger	10	1.058	0.095	2.881 (0.964–8.615)	0.760	–	–
Weighted median	10	0.468	0.132	1.077 (0.869–2.936)	–	–	–
Simple mode	10	0.412	0.412	1.350 (0.591–3.856)	–	–	–
Weighted mode	10	0.429	0.308	1.206 (0.706–3.340)	–	–	–
BWMR	10	0.294	0.212	1.342 (0.846–2.130)	–	–	–
Ruminococcaceae UCG011							
IVW	8	–0.289	0.034	0.749 (0.574–0.978)	0.987	0.485	0.973
MR Egger	8	–0.785	0.293	0.456 (0.120–1.732)	0.983	–	–
Weighted median	8	–0.295	0.080	0.745 (0.535–1.036)	–	–	–
Simple mode	8	–0.344	0.216	0.709 (0.432–1.163)	–	–	–
Weighted mode	8	–0.364	0.222	0.695 (0.408–1.184)	–	–	–
BWMR	8	–0.291	0.04	0.747 (0.565–0.987)	–	–	–

SNPs – single nucleotide polymorphisms; OR – odds ratio; 95% CI – 95% confidence interval; IVW – inverse variance weighted; BWMR – Bayesian weighted Mendelian randomization.

0.100–0.650; p = 0.004) maintained its inverse association with DLBCL risk, whereas 7-HOCA (OR = 10.577, 95% CI: 1.275–87.729; p = 0.029) demonstrated an even stronger positive association. However, 3-(4-hydroxyphenyl)lactate (OR = 0.230, 95% CI: 0.040–1.314; p = 0.098) did not retain statistical significance in the BWMR analysis and was therefore not considered in further interpretation (Table 2). Scatter plots depicting the significant metabolite–DLBCL associations identified through BWMR analysis are presented in Fig. 2D–E. Forest plots visually summarize all significant microbiota–DLBCL causal relationships (Supplementary Fig. 1). Funnel plots demonstrated symmetrical distributions of causal estimates, indicating minimal evidence of directional pleiotropy that could bias the results (Supplementary Fig. 2).

Sensitivity analysis

Sensitivity analyses consistently supported the robustness of our findings. Cochran's Q test revealed no significant heterogeneity among the selected IVs (Tables 1,2). Additionally, the MR-Egger intercept test provided no evidence of significant horizontal pleiotropy. The MR-PRESSO global test (p > 0.05) identified no outlier variants among the gut microbial taxa, microbiota-derived metabolites, or DLBCL associations, further supporting the absence of horizontal pleiotropy (Tables 1,2). Finally, the leave-one-out analysis confirmed the stability of the estimates, demonstrating that no individual variant exerted a disproportionate influence on the causal associations (Supplementary Fig. 3).

Table 2. The causal relationship estimation between gut microbiota-derived metabolites and diffuse large B-cell lymphoma (DLBCL)

Metabolite (exposure)	Number of SNPs	β	p-value	OR (95% CI)	p-value for heterogeneity test	p-value for intercept	p-value for MR-PRESSO
3-(4-hydroxyphenyl)lactate							
IVW	22	−1.749	0.023	0.174 (0.030–0.784)	0.601	0.764	0.561
MR Egger	22	−2.269	0.133	0.103 (0.003–4.043)	0.544	–	–
Weighted median	22	−1.67	0.017	0.188 (0.021–1.682)	–	–	–
Simple mode	22	−2.601	0.188	0.074 (0.001–4.534)	–	–	–
Weighted mode	22	−3.011	0.2	0.049 (0.001–3.616)	–	–	–
BWMR	22	−1.47	0.098	0.23 (0.04–1.314)	–	–	–
7-HOCA							
IVW	15	2.258	0.02	9.561 (1.426–64.088)	0.423	0.561	0.455
MR Egger	15	3.772	0.190	43.47 (0.208–9140.125)	0.375	–	–
Weighted median	15	1.807	0.189	6.092 (0.411–90.161)	–	–	–
Simple mode	15	0.616	0.790	1.851 (0.022–159.266)	–	–	–
Weighted mode	15	1.527	0.484	4.604 (0.071–296.764)	–	–	–
BWMR	15	2.359	0.029	10.577 (1.275–87.729)	–	–	–
Indolepropionate							
IVW	16	−1.217	0.003	0.296 (0.131–0.669)	0.932	0.729	0.914
MR Egger	16	−0.958	0.273	0.384 (0.074–1.990)	0.906	–	–
Weighted median	16	−1.05	0.104	0.350 (0.099–1.239)	–	–	–
Simple mode	16	−1.815	0.116	0.163 (0.019–1.378)	–	–	–
Weighted mode	16	−0.86	0.235	0.423 (0.108–1.654)	–	–	–
BWMR	16	−1.364	0.004	0.256 (0.1–0.65)	–	–	–

SNPs – single nucleotide polymorphisms; OR – odds ratio; 95% CI – 95% confidence interval; IVW – inverse variance weighted; BWMR – Bayesian weighted Mendelian randomization; 7-HOCA – 7α-hydroxy-3-oxo- 4-cholestenoate.

Reverse MR results

In the reverse MR analysis evaluating DLBCL as exposure and gut microbiota as outcome, we initially identified 12 IVs, excluding those with weak instrument effects. Using IVW as the primary statistical method, DLBCL was found to be associated with a higher abundance of *Anaerofilum* (OR = 1.087, 95% CI: 1.008–1.173, p = 0.031) and a lower abundance of Deltaproteobacteria (OR = 0.959, 95% CI: 0.922–0.997, p = 0.037), *Desulfovibrionales* (OR = 0.959, 95% CI: 0.922–0.998, p = 0.041), Desulfovibrionaceae (OR = 0.960, 95% CI: 0.923–0.999, p = 0.045), Oxalobacteraceae (OR = 0.914, 95% CI: 0.843–0.992, p = 0.031), and *Oxalobacter* (OR = 0.909, 95% CI: 0.837–0.988, p = 0.024) (Table 3). The BWMR analysis provided further validation of the causal relationships between DLBCL and specific gut microbial taxa. This robust approach confirmed significant associations of DLBCL with Deltaproteobacteria (OR = 0.958, 95% CI: 0.920–0.999, p = 0.042), *Desulfovibrionales* (OR = 0.959, 95% CI: 0.920–0.999, p = 0.046), Oxalobacteraceae (OR = 0.911, 95% CI: 0.838–0.990, p = 0.027), *Anaerofilum* (OR = 1.091, 95% CI: 1.008–1.181, p = 0.030), and *Oxalobacter* (OR = 0.905, 95% CI: 0.832–0.986, p = 0.022). However, the association with Desulfovibrionaceae (OR = 0.960, 95% CI: 0.921–1.001, p = 0.051) narrowly missed the statistical significance threshold and was consequently excluded from final interpretation. Scatter plots of SNP effects supported the causal associations between DLBCL risk and gut microbiota abundance, as shown in Supplementary Fig. 4.

Forest plots demonstrated significant effects of DLBCL on specific gut microbial taxa in the reverse MR analysis (Supplementary Fig. 5). Funnel plots derived from the reverse MR analysis demonstrated a symmetrical distribution of variant effects, suggesting a low likelihood of directional pleiotropy (Supplementary Fig. 6). Importantly, sensitivity analyses performed for the reverse MR findings provided no evidence of significant heterogeneity or horizontal pleiotropy, with all relevant statistics summarized in Table 3. The leave-one-out analysis further confirmed the stability of the estimates, indicating that no single IV exerted a disproportionate influence on the causal effect estimates (Supplementary Fig. 7).

Discussion

In this study, we applied MR to systematically investigate the causal relationships between gut microbial features, microbiota-derived metabolites, and DLBCL risk.

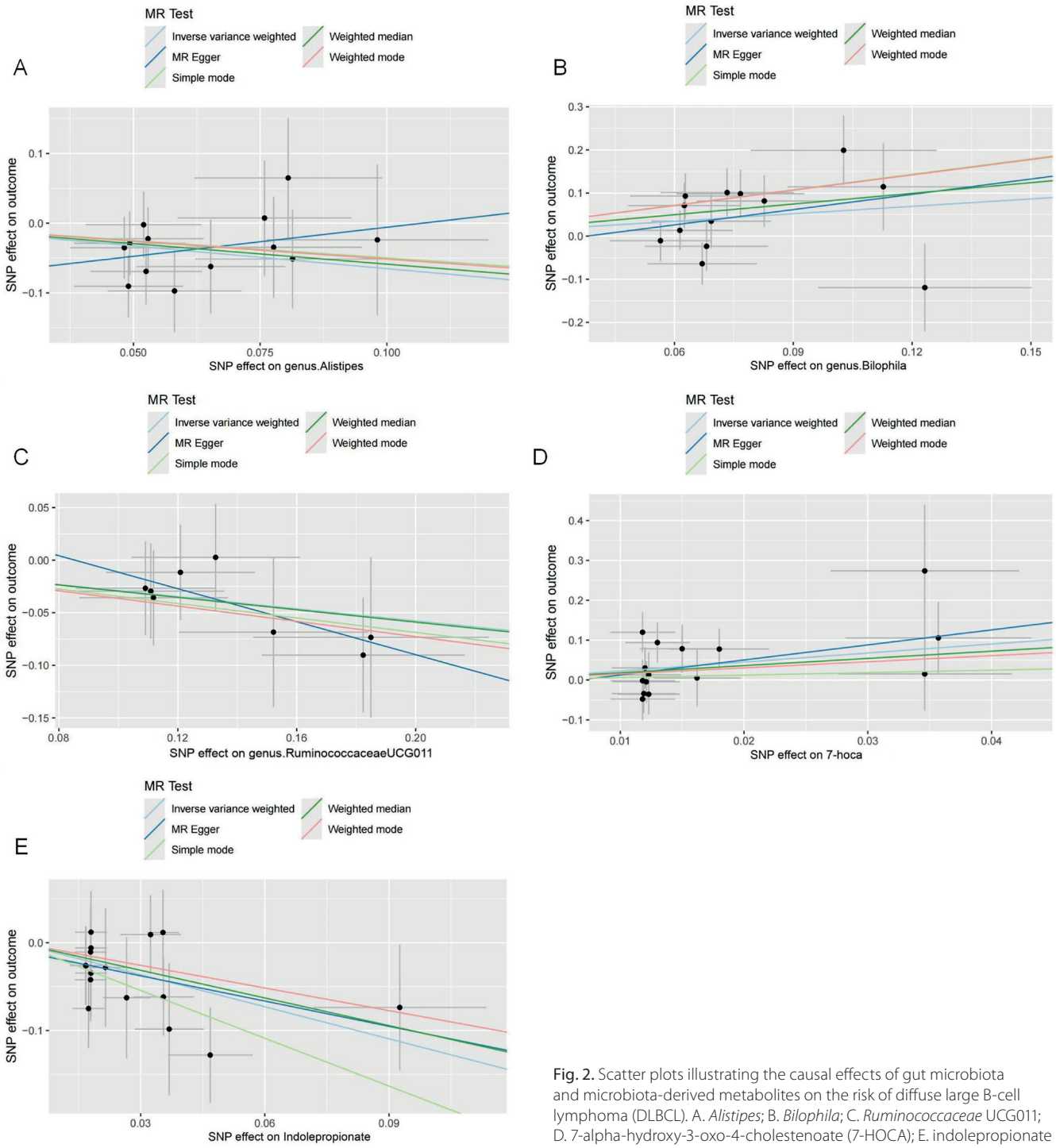

Fig. 2. Scatter plots illustrating the causal effects of gut microbiota and microbiota-derived metabolites on the risk of diffuse large B-cell lymphoma (DLBCL). A. *Alistipes*; B. *Bilophila*; C. *Ruminococcaceae* UCG011; D. 7-alpha-hydroxy-3-oxo-4-cholestenoate (7-HOCA); E. indolepropionate

The findings identified 3 bacterial genera – *Alistipes*, *Ruminococcaceae* UCG011, and *Bilophila* – along with 2 microbiota-derived metabolites, indolepropionate and 7-HOCA, that demonstrated significant causal associations with DLBCL risk. Furthermore, reverse MR analysis suggested that DLBCL may reciprocally influence gut microbial composition, with putative causal effects observed for 1 bacterial class (Deltaproteobacteria), 1 order (*Desulfovibrionales*), 1 family (Oxalobacteraceae), and 2 genera (*Anaerofilum* and *Oxalobacter*). The gut microbiota, comprising more than 1,000 taxonomic units residing in the intestinal tract, is increasingly recognized as a vital "virtual organ" due to its structural complexity and partially heritable characteristics. This microbial community plays a pivotal role in maintaining host physiological homeostasis through diverse direct and indirect mechanisms.[40] Growing evidence suggests that perturbations in gut microbial ecology contribute to the pathogenesis of various diseases, including hematological malignancies.[41,42] Recent studies have further implicated gut microbiota dysbiosis in lymphomagenesis and disease progression.[43,44] Supporting this connection, Lin et al. reported

Table 3. The causal relationship estimation between diffuse large B-cell lymphoma (DLBCL) and gut microbiota

Gut microbiota (outcome)	Number of SNPs	β	p-value	OR (95% CI)	p-value for heterogeneity test	p-value for intercept	p-value for MR-PRESSO
Deltaproteobacteria							
IVW	7	–0.042	0.037	0.959 (0.922–0.997)	0.889	0.508	0.907
MR Egger	7	–0.118	0.326	0.889 (0.719–1.099)	0.876	–	–
Weighted median	7	–0.054	0.030	0.947 (0.902–0.995)	–	–	–
Simple mode	7	–0.058	0.171	0.944 (0.877–1.015)	–	–	–
Weighted mode	7	–0.058	0.137	0.944 (0.884–1.008)	–	–	–
BWMR	7	0.043	0.042	0.958 (0.920–0.999)	–	–	–
Desulfovibrionales							
IVW	7	–0.041	0.041	0.959 (0.922–0.998)	0.889	0.257	0.8
MR Egger	7	–0.120	0.316	0.887 (0.717–1.096)	0.928	–	–
Weighted median	7	–0.054	0.027	0.947 (0.903–0.994)	–	–	–
Simple mode	7	–0.058	0.170	0.944 (0.877–1.015)	–	–	–
Weighted mode	7	–0.057	0.161	0.945 (0.881–1.013)	–	–	–
BWMR	7	–0.042	0.046	0.959 (0.920–0.999)	–	–	–
Desulfovibrionaceae							
IVW	7	–0.041	0.045	0.960 (0.923–0.999)	0.88	0.506	0.893
MR Egger	7	–0.117	0.330	0.890 (0.720–1.010)	0.865	–	–
Weighted median	7	–0.053	0.038	0.948 (0.903–0.996)	–	–	–
Simple mode	7	–0.057	0.169	0.945 (0.882–1.012)	–	–	–
Weighted mode	7	–0.056	0.163	0.946 (0.886–1.010)	–	–	–
BWMR	7	–0.041	0.051	0.960 (0.921–1.0001)	–	–	–
Oxalobacteraceae							
IVW	7	–0.090	0.031	0.914 (0.843–0.992)	0.188	0.930	0.211
MR Egger	7	–0.068	0.792	0.935 (0.581–1.504)	0.120	–	–
Weighted median	7	–0.069	0.172	0.933 (0.845–1.031)	–	–	–
Simple mode	7	–0.068	0.450	0.934 (0.793–1.101)	–	–	–
Weighted mode	7	–0.129	0.137	0.879 (0.759–1.018)	–	–	–
BWMR	7	–0.093	0.027	0.911 (0.838–0.990)	–	–	–
Anaerofilum							
IVW	7	0.083	0.031	1.087 (1.008–1.173)	0.279	0.940	0.333
MR Egger	7	0.101	0.675	1.106 (0.708–1.727)	0.188	–	–
Weighted median	7	0.120	0.011	1.128 (1.028–1.238)	–	–	–
Simple mode	7	0.143	0.091	1.154 (1.004–1.326)	–	–	–
Weighted mode	7	0.138	0.086	1.147 (1.006–1.308)	–	–	–
BWMR	7	0.087	0.03	1.091 (1.008–1.181)	–	–	–
Oxalobacter							
IVW	7	–0.095	0.024	0.909 (0.837–0.988)	0.222	0.995	0.249
MR Egger	7	–0.093	0.722	0.911 (0.560–1.482)	0.144	–	–
Weighted median	7	–0.061	0.244	0.940 (0.848–1.043)	–	–	–
Simple mode	7	–0.010	0.916	0.990 (0.816–1.200)	–	–	–
Weighted mode	7	–0.008	0.938	0.992 (0.822–1.198)	–	–	–
BWMR	7	–0.099	0.022	0.905 (0.832–0.986)	–	–	–

SNPs – single nucleotide polymorphisms; OR – odds ratio; 95% CI – 95% confidence interval; IVW – inverse variance weighted; BWMR – Bayesian weighted Mendelian randomization.

distinct gut microbiota profiles in untreated DLBCL patients, characterized by significantly elevated abundances of Proteobacteria, *Escherichia–Shigella*, *Roseburia*, and *Alistipes* compared with healthy controls.[18] In contrast, Li et al. observed a significant depletion of *Roseburia* in DLBCL patients.[20] Further clinical observations

indicated that patients experiencing treatment-related adverse events exhibited higher levels of Enterobacteriaceae, whereas those without adverse events showed relative enrichment of Prevotellaceae and Oscillospiraceae.[22] Moreover, a significant difference in Enterobacteriaceae abundance was observed between patients with disease relapse or progression and those in remission.[22] Recent investigations have identified significant microbial shifts in treatment-naïve patients, characterized by increased levels of Bacteroidetes and decreased levels of Firmicutes compared with healthy controls.[43] Firmicutes, particularly members of the families Ruminococcaceae and Lachnospiraceae, are major butyrate producers in the human colon.[45,46] In our study, *Ruminococcaceae* UCG011, a butyrate-producing bacterial genus, was inversely associated with DLBCL risk (OR = 0.749). This protective association suggests potential anti-lymphoma effects that may operate through multiple butyrate-mediated mechanisms. As the primary short-chain fatty acid (SCFA) generated through microbial fermentation of dietary fiber, butyrate serves as a key energy source for colonocytes while also modulating immune responses and enhancing intestinal barrier integrity.[47] Moreover, Wei et al. demonstrated that butyrate exerts potent anti-tumor effects by functioning as a histone deacetylase (HDAC) inhibitor, thereby promoting histone acetylation and inducing apoptosis of malignant cells through epigenetic modulation of gene expression pathways.[24] The anti-tumor effects of butyrate may be particularly relevant in DLBCL, given the characteristic overexpression of HDAC isoforms observed in this malignancy.[48,49] Similar HDAC-mediated oncogenic processes have been reported across multiple malignancies, in which butyrate's HDAC-inhibitory activity promotes tumor-suppressive histone acetylation and apoptosis.[50,51]

Existing evidence also indicates that butyrate regulates key oncogenic pathways including mitochondrial and extrinsic apoptosis, G protein-coupled receptor (GPR41/43/109a) signaling, Wnt signaling, and protein kinase C pathway.[52] Particularly relevant to lymphomagenesis, Lu et al. reported that butyrate-producing *Eubacterium rectale* suppresses tumor necrosis factor (TNF) production and subsequent TLR4/MyD88/NF-κB activation in B cells, which may contribute to reduced lymphoma incidence.[23] Our current analysis did not detect significant associations of *E. rectale* with DLBCL risk, potentially due to limited sample size. However, the robust association with *Ruminococcaceae* UCG011 warrants future investigation into its specific mechanisms of action, particularly its potential to modulate these reported butyrate-sensitive pathways in B-cell malignancies.

Alistipes, a recently identified genus of anaerobic bacteria, predominantly colonizes the human gastrointestinal tract but has also been identified in extraintestinal sites including the brain, bloodstream, and gut periphery.[53] This bacterial genus demonstrates complex disease-modulating properties, with studies reporting both protective and pathogenic associations. Several studies have reported that *Alistipes* exerts protective effects against colitis.[54] In patients with liver cirrhosis, *Alistipes* abundance declines progressively from the compensated to the decompensated stage.[55,56] The onset and severity of liver fibrosis have been associated with reduced *Alistipes* populations. The observed antifibrotic effects may be mediated through *Alistipes* production of propionate and acetate, with the latter exhibiting well-documented anti-inflammatory properties.[57,58] Additionally, emerging evidence indicates complex, context-dependent roles for *Alistipes* in cancer biology. While elevated *Alistipes* abundance has demonstrated protective effects against hepatocellular carcinoma (HCC) progression, potentially mediated through the suppression of hepatic T helper 17 cells,[59] other studies suggest this genus may promote carcinogenesis in colorectal cancer via IL-6/STAT3 pathway activation.[60] Our findings revealed a similarly protective association between *Alistipes* and DLBCL risk, which may be mediated through immunomodulatory mechanisms analogous to those observed in hepatic malignancies. However, the precise pathways underlying *Alistipes*–DLBCL interactions require further mechanistic investigation.

Moreover, our MR analysis revealed a significant positive association between *Bilophila* and the risk of DLBCL. The anaerobic genus *Bilophila* has been implicated in diverse pathological conditions, including abscesses, appendicitis, colitis, and Parkinson's disease.[61] Studies have demonstrated that *Bilophila* generates hydrogen sulfide (H_2S), a gaseous metabolite that significantly compromises intestinal barrier function. This disruption facilitates direct contact between harmful substances or bacteria and epithelial surfaces, thereby impairing immune responses, activating inflammation, and ultimately promoting colorectal tumorigenesis.[62,63] Furthermore, evidence also suggests that *Bilophila* exhibits inhibitory effects on butyrate-producing gut microbiota, simultaneously affecting both microbial balance and host physiology.[64] Given the established protective role of butyrate against DLBCL, it is plausible that *Bilophila* may contribute to DLBCL pathogenesis through butyrate depletion mechanisms. However, the precise nature of this relationship remains to be elucidated, as no prior studies have directly investigated *Bilophila*–DLBCL associations.

The gut microbiota serves as a key biosynthetic organ for circulating metabolites, generating diverse small molecules that systemically regulate host physiology through multiple mechanisms. Emerging evidence links dysregulation of these microbial metabolic pathways to various disease states.[65,66] As previously discussed, elevated levels of butyrate have been associated with reduced lymphoma burden, whereas the loss of butyrate-producing bacteria may diminish these beneficial effects.[24] Moreover, this relationship is further supported by observations in NK/T-cell lymphoma patients, who exhibit significantly reduced levels of both butyrate and its primary producer *Faecalibacterium prausnitzii*. Butyrate is hypothesized to inhibit

tumor progression by activating SOCS1 and suppressing the JAK–STAT signaling pathway.[67] Building upon our initial findings, we conducted a comprehensive investigation of the causal relationships between gut microbiota-derived metabolites and DLBCL risk. Our analysis identified indolepropionate, a microbial metabolite generated from dietary tryptophan catabolism,[68] as a potentially protective factor for DLBCL. This compound demonstrates multifaceted biological activity, including the stimulation of interleukin (IL)-10 secretion from bone marrow-derived macrophages to exert potent anti-inflammatory effects.[69] Another study demonstrated that indolepropionate attenuates intestinal inflammation by suppressing interferon gamma (IFN-γ), TNF-α, and IL-1β through its action as an aryl hydrocarbon receptor (AHR) ligand that promotes IL-22 production.[68] Indolepropionate has also been reported to exhibit anti-tumor effects by inhibiting epithelial-mesenchymal transition, enhancing anti-tumor immunity through upregulation of both AHR and pregnane X receptor (PXR) pathways, and functioning as a free radical scavenger against oxidative DNA damage induced by carcinogens like free iron and Cr(III).[70–72] The convergence of these anticancer properties with our MR findings suggests its potential relevance to DLBCL prevention and treatment, although the specific mechanisms in lymphoid malignancies remain to be fully elucidated.

In contrast to well-characterized microbial metabolites, the pathological significance of 7-HOCA remains poorly understood. A recent study reported elevated levels of 7-HOCA in patients with hepatitis and liver cancer and demonstrated that 7-HOCA induces DNA damage and promotes tumorigenesis in non-alcoholic fatty liver disease (NAFLD).[73] In our MR analyses, 7-HOCA was identified as a potential risk factor for DLBCL, providing a compelling rationale for further investigations into the molecular mechanisms through which 7-HOCA may contribute to lymphomagenesis.

This study has several methodological strengths compared with previous research. We employed a MR design to infer causal relationships between gut microbiota, microbiota-derived metabolites, and DLBCL. Bidirectional MR analyses were conducted using large-scale GWAS datasets to comprehensively evaluate these associations. Single nucleotide polymorphisms were selected as instrumental variables to emulate the random allocation of genetic variants. According to Mendel's laws, these variants are randomly assigned at conception and are generally independent of environmental confounders. As highlighted by Ference et al.,[74] this intrinsic property enables MR to estimate causal effects of exposures on disease outcomes with reduced confounding. By serving as proxies for long-term exposure, SNPs facilitate more robust causal inference. To assess potential horizontal pleiotropy, we applied both MR-Egger regression and MR-PRESSO. Finally, the robustness of our findings was further supported through complementary analyses using BWMR.

Limitations of the study

Despite its strengths, this study has several limitations that warrant consideration. First, the GWAS data on gut microbiota and metabolites were predominantly derived from European cohorts, and the DLBCL dataset also consisted exclusively of individuals of European ancestry. Consequently, our findings may be susceptible to population stratification bias and may not be generalizable to non-European populations. Second, the absence of clinical subtyping data precluded important stratified analyses, including the evaluation of differences according to cell-of-origin classification (germinal center B-cell (GCB) vs non-GCB subtypes) or the presence of gastrointestinal involvement. Finally, the resolution of the gut microbiota data was limited to the genus level, and the metabolite dataset was incomplete, thereby constraining more in-depth mechanistic exploration. Future studies incorporating more comprehensive and higher-resolution datasets are warranted to validate and refine the inferred causal relationships between gut microbiota and DLBCL.

Conclusions

In summary, our findings provide evidence supporting potential causal relationships between gut microbiota, microbiota-derived metabolites, and DLBCL, offering new insights into DLBCL pathogenesis and informing future diagnostic and therapeutic strategies. However, given the limitations of this study, further experimental and clinical investigations are warranted to more comprehensively elucidate the roles of gut microbiota and their metabolites in DLBCL development.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.17047566. The package contains the following files:

Supplementary Fig. 1. The forest plots of the causal effects of gut microbiota and derived metabolites on the risk of DLBCL.

Supplementary Fig. 2. The funnel plots of the causal effects of gut microbiota and derived metabolites on the risk of DLBCL.

Supplementary Fig. 3. The leave-one-out analyses of the causal effects of gut microbiota and derived metabolites on the risk of DLBCL.

Supplementary Fig. 4. The scatter plots of the causal effects of DLBCL on the risk of gut microbiota.

Supplementary Fig. 5. The forest plots of DLBCL on the risk of gut microbiota.

Supplementary Fig. 6. The funnel plots of DLBCL on the risk of gut microbiota.

Supplementary Fig. 7. The leave-one-out analyses of DLBCL on the risk of gut microbiota.

Data Availability Statement

All datasets used in this study are publicly available. Gut microbiota GWAS data were obtained from the MiBioGen consortium (https://mibiogen.gcc.rug.nl), metabolite GWAS data from the TwinsUK and KORA studies annotated via the Human Metabolome Database (https://hmdb.ca), and DLBCL summary statistics from the FinnGen consortium (https://www.finngen.fi/en, release R10, December 2023). The code for Mendelian randomization analyses is available in Zenodo at https://doi.org/10.5281/zenodo.17047705.

Consent for publication

Not applicable.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Ganyu Feng https://orcid.org/0009-0000-7063-1025

References

1. Grulich AE, Vajdic CM. The epidemiology of non-Hodgkin lymphoma. *Pathology*. 2005;37(6):409–419. doi:10.1080/00313020500370192
2. Chu Y, Liu Y, Fang X, et al. The epidemiological patterns of non-Hodgkin lymphoma: Global estimates of disease burden, risk factors, and temporal trends. *Front Oncol*. 2023;13:1059914. doi:10.3389/fonc.2023.1059914
3. Li S, Young KH, Medeiros LJ. Diffuse large B-cell lymphoma. *Pathology*. 2018;50(1):74–87. doi:10.1016/j.pathol.2017.09.006
4. Tilly H, Gomes Da Silva M, Vitolo U, et al. Diffuse large B-cell lymphoma (DLBCL): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2015;26:v116–v125. doi:10.1093/annonc/mdv304
5. Martelli M, Ferreri AJM, Agostinelli C, Di Rocco A, Pfreundschuh M, Pileri SA. Diffuse large B-cell lymphoma. *Crit Rev Oncol Hematol*. 2013;87(2):146–171. doi:10.1016/j.critrevonc.2012.12.009
6. Shi Z, Zhang M. Emerging roles for the gut microbiome in lymphoid neoplasms. *Clin Med Insights Oncol*. 2021;15:11795549211024197. doi:10.1177/11795549211024197
7. Upadhyay Banskota S, Skupa SA, El-Gamal D, D'Angelo CR. Defining the role of the gut microbiome in the pathogenesis and treatment of lymphoid malignancies. *Int J Mol Sci*. 2023;24(3):2309. doi:10.3390/ijms24032309
8. Ley RE, Turnbaugh PJ, Klein S, Gordon JI. Human gut microbes associated with obesity. *Nature*. 2006;444(7122):1022–1023. doi:10.1038/4441022a
9. Thursby E, Juge N. Introduction to the human gut microbiota. *Biochem J*. 2017;474(11):1823–1836. doi:10.1042/BCJ20160510
10. Den Besten G, Van Eunen K, Groen AK, Venema K, Reijngoud DJ, Bakker BM. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res*. 2013;54(9):2325–2340. doi:10.1194/jlr.R036012
11. Bäumler AJ, Sperandio V. Interactions between the microbiota and pathogenic bacteria in the gut. *Nature*. 2016;535(7610):85–93. doi:10.1038/nature18849
12. Natividad JMM, Verdu EF. Modulation of intestinal barrier by intestinal microbiota: Pathological and therapeutic implications. *Pharmacol Res*. 2013;69(1):42–51. doi:10.1016/j.phrs.2012.10.007
13. Gensollen T, Iyer SS, Kasper DL, Blumberg RS. How colonization by microbiota in early life shapes the immune system. *Science*. 2016;352(6285):539–544. doi:10.1126/science.aad9378
14. Guevara-Ramírez P, Cadena-Ullauri S, Paz-Cruz E, Tamayo-Trujillo R, Ruiz-Pozo VA, Zambrano AK. Role of the gut microbiota in hematologic cancer. *Front Microbiol*. 2023;14:1185787. doi:10.3389/fmicb.2023.1185787
15. Meng C, Bai C, Brown TD, Hood LE, Tian Q. Human gut microbiota and gastrointestinal cancer. *Genomics Proteomics Bioinformatics*. 2018;16(1):33–49. doi:10.1016/j.gpb.2017.06.002
16. Yu X, Jiang W, Kosik RO, et al. Gut microbiota changes and its potential relations with thyroid carcinoma. *J Adv Res*. 2022;35:61–70. doi:10.1016/j.jare.2021.04.001
17. Zhao Y, Liu Y, Li S, et al. Role of lung and gut microbiota on lung cancer pathogenesis. *J Cancer Res Clin Oncol*. 2021;147(8):2177–2186. doi:10.1007/s00432-021-03644-0
18. Lin Z, Mao D, Jin C, et al. The gut microbiota correlate with the disease characteristics and immune status of patients with untreated diffuse large B-cell lymphoma. *Front Immunol*. 2023;14:1105293. doi:10.3389/fimmu.2023.1105293
19. Nowicka A, Gil L. Microbial metabolomics in acute myeloid leukemia: From pathogenesis to treatment [published online as ahead of print on October 21, 2025]. *Adv Clin Exp Med*. 2024. doi:10.17219/acem/191559
20. Yuan L, Wang W, Zhang W, et al. Gut microbiota in untreated diffuse large B cell lymphoma patients. *Front Microbiol*. 2021;12:646361. doi:10.3389/fmicb.2021.646361
21. Zhang Y, Han S, Xiao X, et al. Integration analysis of tumor metagenome and peripheral immunity data of diffuse large-B cell lymphoma. *Front Immunol*. 2023;14:1146861. doi:10.3389/fimmu.2023.1146861
22. Yoon SE, Kang W, Choi S, et al. The influence of microbial dysbiosis on immunochemotherapy-related efficacy and safety in diffuse large B-cell lymphoma. *Blood*. 2023;141(18):2224–2238. doi:10.1182/blood.2022018831
23. Lu H, Xu X, Fu D, et al. Butyrate-producing *Eubacterium rectale* suppresses lymphomagenesis by alleviating the TNF-induced TLR4/MyD88/NF-κB axis. *Cell Host Microbe*. 2022;30(8):1139–1150.e7. doi:10.1016/j.chom.2022.07.003
24. Wei W, Sun W, Yu S, Yang Y, Ai L. Butyrate production from high-fiber diet protects against lymphoma tumor. *Leuk Lymphoma*. 2016;57(10):2401–2408. doi:10.3109/10428194.2016.1144879
25. Rinninella E, Raoul P, Cintoni M, et al. What is the healthy gut microbiota composition? A changing ecosystem across age, environment, diet, and diseases. *Microorganisms*. 2019;7(1):14. doi:10.3390/microorganisms7010014
26. Emdin CA, Khera AV, Kathiresan S. Mendelian randomization. *JAMA*. 2017;318(19):1925. doi:10.1001/jama.2017.17219
27. Long Y, Tang L, Zhou Y, Zhao S, Zhu H. Causal relationship between gut microbiota and cancers: A two-sample Mendelian randomisation study. *BMC Med*. 2023;21(1):66. doi:10.1186/s12916-023-02761-6
28. Ma J, Li J, Jin C, et al. Association of gut microbiome and primary liver cancer: A two-sample Mendelian randomization and case–control study. *Liver Int*. 2023;43(1):221–233. doi:10.1111/liv.15466
29. Van Der Velde KJ, Imhann F, Charbon B, et al. MOLGENIS research: Advanced bioinformatics data software for non-bioinformaticians. *Bioinformatics*. 2019;35(6):1076–1078. doi:10.1093/bioinformatics/bty742
30. Li P, Wang H, Guo L, et al. Association between gut microbiota and preeclampsia-eclampsia: A two-sample Mendelian randomization study. *BMC Med*. 2022;20(1):443. doi:10.1186/s12916-022-02657-x
31. Pierce BL, Ahsan H, VanderWeele TJ. Power and instrument strength requirements for Mendelian randomization studies using multiple genetic variants. *Int J Epidemiol*. 2011;40(3):740–752. doi:10.1093/ije/dyq151
32. Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. *Genet Epidemiol*. 2013;37(7):658–665. doi:10.1002/gepi.21758
33. Burgess S, Dudbridge F, Thompson SG. Combining information on multiple instrumental variables in Mendelian randomization: Comparison of allele score and summarized data methods. *Statist Med*. 2016;35(11):1880–1906. doi:10.1002/sim.6835
34. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: Effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015;44(2):512–525. doi:10.1093/ije/dyv080

35. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. *Genet Epidemiol*. 2016;40(4):304–314. doi:10.1002/gepi.21965
36. Hemani G, Zheng J, Elsworth B, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife*. 2018;7:e34408. doi:10.7554/eLife.34408
37. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet*. 2018;50(5):693–698. doi:10.1038/s41588-018-0099-7
38. Hemani G, Tilling K, Davey Smith G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLoS Genet*. 2017;13(11):e1007081. doi:10.1371/journal.pgen.1007081
39. Zhao J, Ming J, Hu X, Chen G, Liu J, Yang C. Bayesian weighted Mendelian randomization for causal inference based on summary statistics. *Bioinformatics*. 2020;36(5):1501–1508. doi:10.1093/bioinformatics/btz749
40. Baquero F, Nombela C. The microbiome as a human organ. *Clin Microbiol Infect*. 2012;18:2–4. doi:10.1111/j.1469-0691.2012.03916.x
41. Gagnière J. Gut microbiota imbalance and colorectal cancer. *World J Gastroenterol*. 2016;22(2):501. doi:10.3748/wjg.v22.i2.501
42. Lucas C, Barnich N, Nguyen H. Microbiota, inflammation and colorectal cancer. *Int J Mol Sci*. 2017;18(6):1310. doi:10.3390/ijms18061310
43. Diefenbach CS, Peters BA, Li H, et al. Microbial dysbiosis is associated with aggressive histology and adverse clinical outcome in B-cell non-Hodgkin lymphoma. *Blood Adv*. 2021;5(5):1194–1198. doi:10.1182/bloodadvances.2020003129
44. Yamamoto ML, Maier I, Dang AT, et al. Intestinal bacteria modify lymphoma incidence and latency by affecting systemic inflammatory state, oxidative stress, and leukocyte genotoxicity. *Cancer Res*. 2013;73(14):4222–4232. doi:10.1158/0008-5472.CAN-13-0022
45. Barcenilla A, Pryde SE, Martin JC, et al. Phylogenetic relationships of butyrate-producing bacteria from the human gut. *Appl Environ Microbiol*. 2000;66(4):1654–1661. doi:10.1128/AEM.66.4.1654-1661.2000
46. Louis P, Duncan SH, McCrae SI, Millar J, Jackson MS, Flint HJ. Restricted distribution of the butyrate kinase pathway among butyrate-producing bacteria from the human colon. *J Bacteriol*. 2004;186(7):2099–2106. doi:10.1128/JB.186.7.2099-2106.2004
47. Liu H, Wang J, He T, et al. Butyrate: A double-edged sword for health? *Adv Nutr*. 2018;9(1):21–29. doi:10.1093/advances/nmx009
48. Lee SH, Yoo C, Im S, Jung JH, Choi HJ, Yoo J. Expression of histone deacetylases in diffuse large B-cell lymphoma and its clinical significance. *Int J Med Sci*. 2014;11(10):994–1000. doi:10.7150/ijms.8522
49. Johnson DP, Spitz GS, Tharkar S, et al. HDAC1,2 inhibition impairs EZH2- and BBAP-mediated DNA repair to overcome chemoresistance in EZH2 gain-of-function mutant diffuse large B-cell lymphoma. *Oncotarget*. 2015;6(7):4863–4887. doi:10.18632/oncotarget.3120
50. Donohoe DR, Holley D, Collins LB, et al. A gnotobiotic mouse model demonstrates that dietary fiber protects against colorectal tumorigenesis in a microbiota- and butyrate-dependent manner. *Cancer Discov*. 2014;4(12):1387–1397. doi:10.1158/2159-8290.CD-14-0501
51. Stojanovic N, Hassan Z, Wirth M, et al. HDAC1 and HDAC2 integrate the expression of p53 mutants in pancreatic cancer. *Oncogene*. 2017;36(13):1804–1815. doi:10.1038/onc.2016.344
52. Chen J, Zhao KN, Vitetta L. Effects of intestinal microbial-elaborated butyrate on oncogenic signaling pathways. *Nutrients*. 2019;11(5):1026. doi:10.3390/nu11051026
53. Shkoporov AN, Chaplin AV, Khokhlova EV, et al. *Alistipes inops* sp. nov. and *Coprobacter secundus* sp. nov., isolated from human faeces. *Int J Syst Evol Microbiol*. 2015;65(Pt 12):4580–4588. doi:10.1099/ijsem.0.000617
54. Dziarski R, Park SY, Kashyap DR, Dowd SE, Gupta D. Pglyrp-regulated gut microflora *Prevotella falsenii*, *Parabacteroides distasonis* and *Bacteroides eggerthii* enhance and *Alistipes finegoldii* attenuates colitis in mice. *PLoS One*. 2016;11(1):e0146162. doi:10.1371/journal.pone.0146162
55. Iebba V, Guerrieri F, Di Gregorio V, et al. Combining amplicon sequencing and metabolomics in cirrhotic patients highlights distinctive microbiota features involved in bacterial translocation, systemic inflammation and hepatic encephalopathy. *Sci Rep*. 2018;8(1):8210. doi:10.1038/s41598-018-26509-y
56. Shao L, Ling Z, Chen D, Liu Y, Yang F, Li L. Disorganized gut microbiome contributed to liver cirrhosis progression: A meta-omics-based study. *Front Microbiol*. 2018;9:3166. doi:10.3389/fmicb.2018.03166
57. Oliphant K, Allen-Vercoe E. Macronutrient metabolism by the human gut microbiome: Major fermentation by-products and their impact on host health. *Microbiome*. 2019;7(1):91. doi:10.1186/s40168-019-0704-8
58. Parker BJ, Wearsch PA, Veloo ACM, Rodriguez-Palacios A. The genus *Alistipes*: Gut bacteria with emerging implications to inflammation, cancer, and mental health. *Front Immunol*. 2020;11:906. doi:10.3389/fimmu.2020.00906
59. Li J, Sung CYJ, Lee N, et al. Probiotics modulated gut microbiota suppresses hepatocellular carcinoma growth in mice. *Proc Natl Acad Sci U S A*. 2016;113(9):E1306-15. doi:10.1073/pnas.1518189113
60. Moschen AR, Gerner RR, Wang J, et al. Lipocalin 2 protects from inflammation and tumorigenesis associated with gut microbiota alterations. *Cell Host Microbe*. 2016;19(4):455–469. doi:10.1016/j.chom.2016.03.007
61. Burrichter AG, Dörr S, Bergmann P, et al. Bacterial microcompartments for isethionate desulfonation in the taurine-degrading human-gut bacterium *Bilophila wadsworthia*. *BMC Microbiol*. 2021;21(1):340. doi:10.1186/s12866-021-02386-w
62. Waqas M, Halim SA, Ullah A, et al. Multi-fold computational analysis to discover novel putative inhibitors of isethionate sulfite-lyase (Isla) from *Bilophila wadsworthia*: Combating colorectal cancer and inflammatory bowel diseases. *Cancers (Basel)*. 2023;15(3):901. doi:10.3390/cancers15030901
63. Yazici C, Wolf PG, Kim H, et al. Race-dependent association of sulfidogenic bacteria with colorectal cancer. *Gut*. 2017;66(11):1983–1994. doi:10.1136/gutjnl-2016-313321
64. Natividad JM, Lamas B, Pham HP, et al. Bilophila wadsworthia aggravates high fat diet induced metabolic dysfunctions in mice. *Nat Commun*. 2018;9(1):2802. doi:10.1038/s41467-018-05249-7
65. Yue X, Zhou H, Wang S, Chen X, Xiao H. Gut microbiota, microbiota-derived metabolites, and graft-versus-host disease. *Cancer Med*. 2024;13(3):e6799. doi:10.1002/cam4.6799
66. D'Angelo CR, Sudakaran S, Callander NS. Clinical effects and applications of the gut microbiome in hematologic malignancies. *Cancer*. 2021;127(5):679–687. doi:10.1002/cncr.33400
67. Shi Z, Li M, Zhang C, et al. Butyrate-producing *Faecalibacterium prausnitzii* suppresses natural killer/T-cell lymphoma by dampening the JAK-STAT pathway. *Gut*. 2025;74(4):557–570. doi:10.1136/gutjnl-2024-333530
68. Jiang H, Chen C, Gao J. Extensive summary of the important roles of indole propionic acid, a gut microbial metabolite in host health and disease. *Nutrients*. 2022;15(1):151. doi:10.3390/nu15010151
69. Wlodarska M, Luo C, Kolde R, et al. Indoleacrylic acid produced by commensal *Peptostreptococcus* species suppresses inflammation. *Cell Host Microbe*. 2017;22(1):25–37.e6. doi:10.1016/j.chom.2017.06.007
70. Sári Z, Mikó E, Kovács T, et al. Indolepropionic acid, a metabolite of the microbiome, has cytostatic properties in breast cancer by activating AHR and PXR receptors and inducing oxidative stress. *Cancers (Basel)*. 2020;12(9):2411. doi:10.3390/cancers12092411
71. Qi W, Reiter RJ, Tan D, Manchester LC, Siu AW, Garcia JJ. Increased levels of oxidatively damaged DNA induced by chromium(III) and H_2O_2: protection by melatonin and related molecules. *J Pineal Res*. 2000;29(1):54–61. doi:10.1034/j.1600-079X.2000.290108.x
72. Karbownik M, Reiter RJ, Garcia JJ, et al. Indole-3-propionic acid, a melatonin-related molecule, protects hepatic microsomal membranes from iron-induced oxidative damage: Relevance to cancer reduction. *J Cell Biochem*. 2001;81(3):507–513. PMID:11255233.
73. Nikolaou N, Arvaniti A, Sanna F, et al. AKR1D1 knockdown identifies 7[alpha]-hydroxy-3-oxo-4-cholestenoic acid (7-HOCA) as a driver of metabolic dysfunction and hepatocellular cancer risk in patients with non-alcoholic fatty liver disease (NAFLD). *Endocr Abstracts*. 2022;81:Y18. doi:10.1530/endoabs.81.YI8
74. Ference BA, Holmes MV, Smith GD. Using Mendelian randomization to improve the design of randomized trials. *Cold Spring Harb Perspect Med*. 2021;11(7):a040980. doi:10.1101/cshperspect.a040980

Brain volumetric variability and artificial intelligence diagnosis: Importance of race/ethnicity-specific reference standards and social determinant adjustment. A scoping review

Srinivasa Rao Bolla[1,A–D,F], Joseph Uy Almazan[2,A,C,E,F], Rauan Satbekov[3,A–D,F], Syed Hani Abidi[1,E,F], Dinara Jumadilova[2,E,F], Kamila Mussabekova[4,E,F], Surendra Maharjan[5,E,F]

[1] Department of Biomedical Sciences, School of Medicine, Nazarbayev University, Astana, Kazakhstan
[2] Department of Medicine, School of Medicine, Nazarbayev University, Astana, Kazakhstan
[3] Department of Surgery, School of Medicine, Nazarbayev University, Astana, Kazakhstan
[4] Consultation and Diagnostic Department, National Center of Neurosurgery, Astana, Kazakhstan
[5] Department of Radiology, Weill Cornell Medical College, New York, 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(5):889–903

Address for correspondence
Srinivasa Rao Bolla
E-mail: srinivasa.bolla@nu.edu.kz

Funding sources
This work was supported by Nazarbayev University under the Faculty Development Competitive Research Grants Program (2024–2026), grant No. 201223FD2605-AI (principal investigator: Srinivasa Rao Bolla).

Conflict of interest
None declared

Received on September 8, 2024
Reviewed on June 11, 2025
Accepted on July 31, 2025

Published online on March 30, 2026

Cite as
Bolla SR, Almazan JU, Satbekov R, et al. Brain volumetric variability and artificial intelligence diagnosis: Importance of race/ethnicity-specific reference standards and social determinant adjustment. A scoping review. *Adv Clin Exp Med*. 2026;35(5):889–903. doi:10.17219/acem/208841

DOI
10.17219/acem/208841

Abstract

Neuroimaging techniques such as magnetic resonance imaging (MRI) are routinely used in diagnostic radiology to evaluate brain changes associated with neurological and psychiatric conditions. Evidence suggests that imaging biomarkers predict clinical outcomes with varying accuracy across ethnic groups. Underrepresentation of ethnic diversity in neuroimaging research may result in generalization bias, whereby findings derived from one population are inaccurately applied to others. A scoping review methodology was employed to systematically identify and analyze relevant literature. Searches were conducted across EBSCOhost (including CINAHL and Medline Complete), Elsevier (Scopus), Clarivate (Web of Science), and PubMed. Eligible studies examined ethnicity-related differences in subcortical brain volumes and cortical thickness in healthy adults using neuroimaging. The search yielded 1,013 records, which were screened according to predefined inclusion and exclusion criteria. Fourteen studies met the eligibility criteria and were included in the final analysis. The reviewed studies demonstrated significant variations in cortical thickness and subcortical volumes across diverse populations and socioeconomic groups, underscoring the need for population-sensitive reference standards in neuroimaging to minimize generalization bias. These findings highlight the importance of incorporating ethnic variability into neuroimaging research and developing population-sensitive frameworks for MRI-based diagnostic applications. Additionally, the review identifies key areas for improvement, including the integration of ethnic and socioeconomic diversity in artificial intelligence (AI)-driven neuroimaging models to enhance diagnostic precision and equity.

Key words: brain, magnetic resonance imaging, cortex, artificial intelligence, ethnicity

Highlights

- Significant volumetric differences in cortical and subcortical brain structures are observed across ethnic populations.
- Aging, socioeconomic status (SES), and racialized social identity interact to influence brain anatomy, cognitive performance, and health outcomes.
- Population-sensitive neuroimaging templates incorporating race/ethnicity and SES can improve AI-based diagnostic accuracy for neurodegenerative diseases.
- Ethical AI development in neuroimaging requires bias mitigation, data privacy protection, and healthcare accountability.

Introduction

Alzheimer's and Parkinson's diseases are among the most prevalent neurodegenerative disorders, significantly impacting global health. Over 57 million people were living with dementia worldwide in 2019.[1] This number is projected to rise to more than 153 million by 2050 due to population aging. Neuroimaging techniques, particularly magnetic resonance imaging (MRI) and positron emission tomography (PET), are invaluable for diagnosis. The integration of artificial intelligence (AI) into imaging enhances segmentation, lesion detection, and volumetric analysis, thereby facilitating the identification of subtle structural changes and improving early diagnosis of neurodegenerative and psychiatric disorders.[2]

Beyond neuroimaging, AI applications in health sciences continue to expand. For example, Razdan et al.[3] explored the role of AI in pediatric dentistry by analyzing dentists' perceptions of its diagnostic capabilities. Furthermore, AI models have been developed to assess pediatric frailty syndrome in heart failure[4] and pediatric-onset multiple sclerosis.[5] The cortical and subcortical brain regions are involved in higher-order cognitive processes, including decision-making, language, complex motor control, emotion, motivation, and basic survival functions, and play a fundamental role in behavior. Specific cognitive domains are associated with distinct brain structures. Memory is linked to hippocampal and entorhinal cortex volumes; attention and processing speed are related to parietal lobe volumes and white matter integrity[6]; executive functions involve the prefrontal cortex and caudate[7,8]; emotional processing and social cognition are associated with amygdala volume[8]; and global cognitive and social functioning have been linked to thalamic volume.[9] Additionally, the prefrontal cortex (PFC) plays a key role in executive functioning, emotional regulation, and decision-making,[10] while the pre-supplementary motor cortex and the right inferior frontal gyrus are implicated in reactive action control.[11] These brain structures serve as important biomarkers of neurological disorders; their morphology aids diagnosis, monitoring of disease progression, and prognostication. Alterations in cortical thickness and regional brain volumes – including total brain volume, parietal and temporal gray matter, and the hippocampus – are well-established biomarkers of neurodegenerative diseases such as Alzheimer's disease (AD) and related dementias,[12–14] Parkinson's disease (PD),[15] Huntington's disease,[16] amyotrophic lateral sclerosis (ALS),[17] and global cognitive decline.[6] Cortical structural abnormalities are likewise observed in psychiatric disorders, including major depressive disorder,[18,19] schizophrenia,[20] and bipolar disorder.[10,21,22]

Furthermore, volumetric measures of subcortical structures – including the thalamus, basal ganglia, hippocampus, amygdala, and nucleus accumbens – serve as key diagnostic biomarkers. Subcortical structural alterations have been documented in a range of neurological and psychiatric disorders, including AD,[23] PD,[24] schizophrenia, bipolar disorder, major depressive disorder, and autism spectrum disorder.[25–27]

Ethnic differences in brain volumes in both health and disease are well documented in the literature and have important implications for clinical practice and research. For example, studies comparing Caucasian and African American populations,[28] as well as Korean and other East Asian cohorts,[29,30] have reported significant differences in cortical and subcortical volumes. In a multiethnic study, Black participants demonstrated larger total brain volumes, while both Black and Hispanic participants exhibited higher white matter volumes compared with other groups.[31] Differences in pathological markers have also been observed. Black participants showed a greater burden of white matter hyperintensities (WMH), whereas Japanese individuals demonstrated less functional disability despite comparable levels of brain atrophy.[31]

Understanding differences in these brain biomarkers is crucial for accurate diagnosis, particularly in the context of AI-driven neuroimaging. Although genetic ancestry may contribute to intergroup variation in brain morphology,[30] such differences are more consistently associated with social determinants of health, including income, education, access to healthcare, and chronic stress exposure.[31,32] Recognizing race as a sociopolitical construct rather than a purely biological determinant[33,34] helps prevent the perpetuation of biological essentialism within AI-based diagnostic frameworks.[35,36] A significant challenge arises from the fact that commercially available AI-enabled neuroimaging software is primarily trained on data derived from

geographically limited populations, particularly in the USA, Europe, and China,[37–40] thereby introducing potential bias. Consequently, the development of inclusive AI models with broader population representation – integrating biological, social, and geographic determinants – is essential to enhance diagnostic equity and generalizability.

This review critically examines evidence from the literature on how racial, ethnic, and social determinants, conceptualized as sociopolitical and structural factors, influence brain volumetric variability. Throughout the review, racial and ethnic groupings are used within a socio-structural context rather than as biologically fixed categories. Moreover, we adopt the term "population-sensitive" to describe approaches that consider both biological variation (e.g., ancestry and sex) and socio-structural determinants (e.g., race, ethnicity, and socioeconomic status).

Although the cited literature variably uses the terms "race," "ethnicity," and related constructs, our use of the term "population-sensitive" aims to provide a more inclusive and precise framework that avoids reinforcing essentialist or reductive interpretations. While acknowledging the diagnostic benefits of race-aware AI models and population-specific templates, we also highlight the associated ethical risks, including racial essentialism and algorithmic bias. This review emphasizes the importance of evidence-based, bias-reducing strategies that prioritize diagnostic accuracy, equity, and scientific rigor.

Objectives

This scoping review explores differences in cortical and subcortical volumes among diverse populations and addresses the following research questions:

– What is known about cerebral cortical and subcortical volumetric differences across ethnic groups, and how do these differences affect AI-based neuroimaging tools?

– What is the role of specific cortical and subcortical brain structures in health and in neurological and psychiatric conditions?

– How do normative brain volumes across different population groups influence diagnostic accuracy, AI model generalization, and population-sensitive standards?

– What are the potential challenges, biases, and ethical considerations involved in incorporating ethnicity and socioeconomic status (SES) when developing population-sensitive norms and AI models in neuroimaging?

Concept

This scoping review, based on MRI studies, examines population-specific differences in brain volumetry among healthy adults and their implications for ensuring accurate diagnoses and equitable AI-driven assessments. As described earlier, the term "population-sensitive" is used to highlight a comprehensive approach that integrates biological variation (e.g., ancestry and sex) with socio-structural influences (e.g., race, ethnicity, and socioeconomic status). A population-sensitive perspective emphasizes diagnostic precision and equity while preventing the reinforcement of racial essentialism in AI development.

Context

Given the increasing integration of AI into diagnostic processes, population-level diversity in brain volumetry is a key factor in interpreting MRI scans, necessitating population-sensitive reference standards. This review identifies key themes and evaluates the feasibility of personalized diagnostic standards to enhance the accuracy and fairness of AI-based neuroimaging.

Materials and methods

This scoping review follows the framework proposed by Arksey and O'Malley[41] and later refined by Levac et al.[42] to systematically explore ethnic differences in brain volumetry among healthy adults using MRI data. Scoping reviews are useful for mapping key concepts in emerging fields, identifying research gaps, and guiding future investigations.[43] Although scoping reviews do not require formal quality appraisal,[44] this study enhances transparency and reliability by acknowledging the methodological limitations of the included studies and presenting a structured summary table.

Search strategy

Initially, we conducted a preliminary search across the EBSCOhost databases (including CINAHL (Cumulative Index to Nursing and Allied Health Literature) and MEDLINE Complete), Elsevier databases (Scopus), Clarivate databases (Web of Science), and PubMed, which yielded no published reviews on the specified topic. Subsequently, a 2nd search was performed across the same databases to identify relevant articles published between 2014 and 2024, in order to capture the most recent data and refine the search output. The search terms were ("brain anatomy" OR "brain size" OR "brain structure" OR "brain morphology" OR "brain volume" OR "brain volumetric" OR "brain morphometry") AND ("race" OR "racial" OR "ethnicity" OR "ethnic" OR "racioethnic").

To enhance clarity and provide specific support for the need for population-sensitive diagnostic standards, the search specifically focused on studies examining ethnic variability in brain morphology. We included studies comparing cortical and subcortical brain volumes across different racial and ethnic groups to ensure that our review captured the full scope of morphological differences relevant to neuroimaging practice.

This review included studies that analyzed brain volumes using MRI across different ethnic groups, as well

as studies that compared brain volumes among different ethnicities. Only studies involving healthy adult populations were included. The search encompassed peer-reviewed studies published in English and other languages.

Studies were excluded if they: focused primarily on adolescent or pediatric populations; did not compare volumetric outcomes across racial or ethnic groups; used imaging modalities other than MRI; or lacked full-text availability.

Selection of studies

All citations retrieved through the database search were imported into the Zotero reference management software (https://www.zotero.org). Two independent reviewers (A.J. and S.R.) assessed the titles and abstracts to determine their relevance according to the inclusion criteria. Disagreements regarding study inclusion were resolved through discussion to ensure consistency in the selection process. Following the initial screening phase, full-text versions of the selected articles were reviewed, and the exclusion criteria were applied. Excluded studies were documented, with justifications provided for their removal. The final selection process followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) flowchart (Fig. 1). A standardized data extraction form was used to chart study design, population characteristics, MRI modalities, segmentation software, and key findings. Two reviewers (A.J. and S.R.) independently extracted the data, and discrepancies were resolved by consensus.

Study heterogeneity

Studies included in this review varied in MRI acquisition protocols, analytical approaches, and population demographics, which may introduce inconsistencies in the findings. To manage this heterogeneity, we: 1) identified common study designs (predominantly cross-sectional); 2) highlighted differences in MRI modalities and processing methods; and 3) evaluated key limitations, such as scanner variability, small sample sizes, and confounder adjustment (e.g., for age, sex, and SES).

Quality assessment and methodological limitations

Although a formal risk-of-bias assessment was not mandated, we conducted a structured methodological appraisal using the Joanna Briggs Institute (JBI) Checklist for Analytical Cross-Sectional and Cohort Studies.[44] The identified limitations are summarized in Table 1[28–30,45–55] and include: variability in MRI protocols; differences in volumetric measurement software; limited ethnic diversity and generalizability; and inconsistent adjustment for

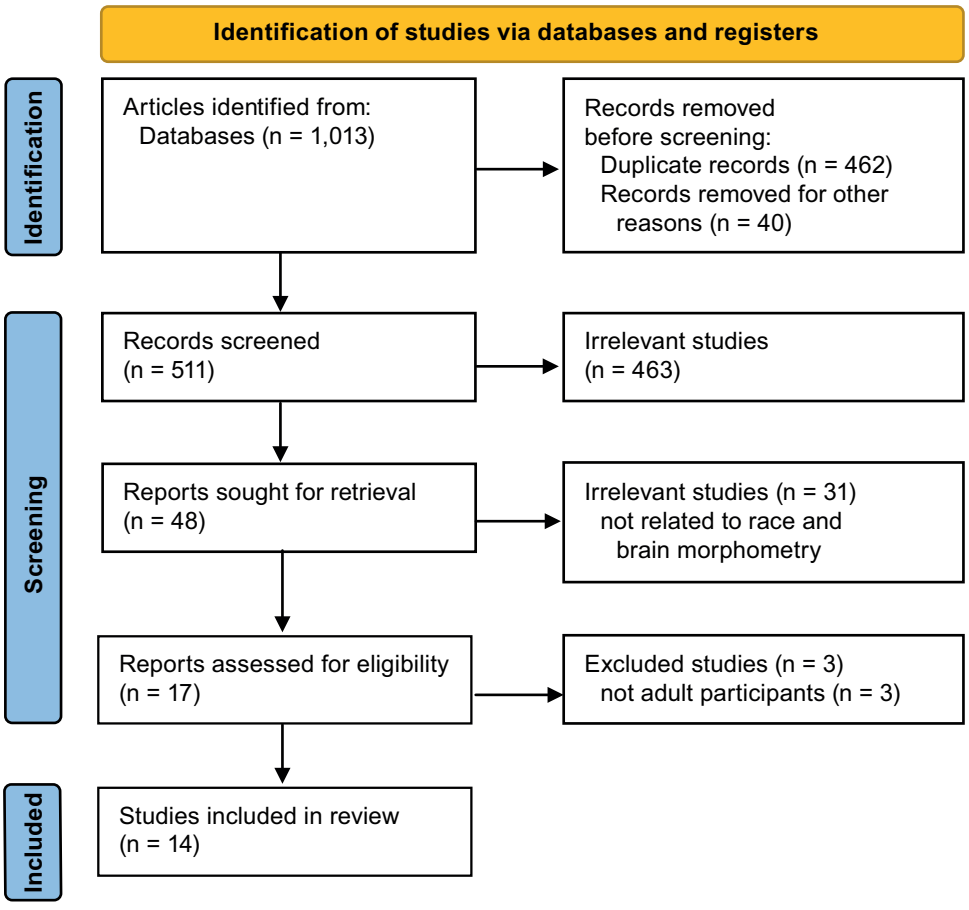

Fig. 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) flowchart of the study selection process

Table 1. Summary of characteristics of the studies included in this review

Study	Authors (year), country	Study design	Population, context	Analysis	Limitations	Key findings
1	Atilano-Barbosa and Barrios (2023),[28] Mexico	cross-sectional observational study	White (n = 338) or African American (n = 56) participants	Nonparametrical permutation analysis of covariance (ANCOVA).	- Unbalanced sample with overrepresentation of White participants. - Confounding variables in the Human Connectome Project (HCP) dataset. - Self-reported race without genetic ancestry data. - Restricted statistical power due to imbalance.	- Morphological brain differences were observed between African American and White participants across white matter and forebrain, midbrain, and hindbrain regions.
2	Choi et al. (2021),[29] South Korea	cross-sectional observational study	Korean (n = 1,686) and Caucasian (n = 851) participants	Regression analysis to predict brain volumes; logistic regression modeling for Alzheimer's disease (AD) diagnosis using regional z-scores; propensity score matching for validation groups; and area under the curve (AUC) comparisons using bootstrap analysis.	- Use of multiple magnetic resonance imaging (MRI) scanners for the Caucasian sample, potentially increasing variance in image quality and measurements. - Cross-sectional study design limiting assessment of individual-level longitudinal changes. - Inherent limitations of longitudinal studies (e.g., small sample sizes, limited follow-up duration).	- Ethnicity was significantly associated with volumetric differences across all lobar and subcortical regions, except for the left pallidum and bilateral ventricles. - Brain volume z-scores derived from a prediction model incorporating ethnicity as a predictor improved diagnostic accuracy for multi-ethnic patients with AD.
3	Kang et al. (2020),[30] South Korea	cross-sectional observational study	East Asian (n = 171) and Caucasian (n = 178) participants	Descriptive statistics; group comparison tests (two-sample t-tests and χ^2 tests); regression analysis (general linear model); vertex-wise statistical analysis; and correction for multiple comparisons.	- Scanner heterogeneity across centers, potentially reducing generalizability. - Slight differences in voxel size between groups (0.9 mm vs 1.0–1.25 mm), potentially introducing measurement bias. - Inference of functional implications (e.g., language or cognition) from structural data without direct functional assessment. - East Asian sample limited to Korean participants, restricting generalizability to other East Asian populations (e.g., Chinese or Japanese). - Statistically significant difference in Mini-Mental State Examination (MMSE) scores between groups despite both being cognitively normal (0.4-point difference).	- Differences in cortical anatomy between cognitively normal East Asian and Caucasian older adults were observed.
4	Lou et al. (2019),[45] China	cross-sectional observational study	Chinese (n = 45) and Caucasian (n = 45) participants	Parametric tests (t-test, analysis of variance (ANOVA)), linear regression, correction for multiple comparisons, and Vertex-wise surface-based morphometric analyses.	- Sample restricted to right-handed male participants, precluding assessment of sex and handedness effects on brain asymmetry. - Relatively small and demographically homogeneous sample, limiting generalizability. - Absence of behavioral and functional magnetic resonance imaging (fMRI) data, limiting interpretation of structural asymmetries in relation to language and cognition.	- Interhemispheric differences were observed between Chinese and Caucasian participants. - Chinese participants exhibited greater structural asymmetries in the frontal, temporal, occipital, and insular cortices compared with Caucasian participants. - These structural differences may be associated with functional disparities between the 2 populations, with potential implications for language learning and visuospatial processing.

Table 1. Summary of characteristics of the studies included in this review – cont.

Study	Authors (year), country	Study design	Population, context	Analysis	Limitations	Key findings
5	Lu et al. (2016),[46] China	comparative cross-sectional study	Uyghurs (n = 15) and Han Chinese (n = 15) participants	Vertex-wise cortical surface group analysis using a 2-class general linear model (GLM), random effects analysis, and statistical parametric mapping, normality tests were performed, and double sample t-tests were used for group comparisons.	- Lack of population-specific cognitive and behavioral data for the Uyghur group, limiting correlation of MRI findings with cognitive or social behavioral differences. - Hypothesized language-driven anatomical plasticity without functional imaging (e.g., fMRI) to verify causal associations. - Sample restricted to young adults (~18 years), limiting lifespan generalizability and developmental interpretation. - Small sample size (n = 15 per group), potentially reducing statistical power and stability of findings.	- Structural brain differences were observed between young Uyghur and Han Chinese participants, consistent with prior East–West comparisons.
6	Lu et al. (2017),[47] China	cross-sectional observational study	Chinese (n = 36) and Indian (n = 32) participants	Student's t-test, χ^2 test, and Bonferroni correction.	- Restriction to region-of-interest (ROI)-based analysis. - Small and demographically homogeneous sample size. - Limited ethnic diversity and narrow age range.	- Population-related differences in brain morphometry were observed between Chinese and Indian undergraduate participants. - Differences in gray matter volume, cortical thickness, and surface area were identified, with potential implications for cognitive function.
7	Rao et al. (2017),[48] India	comparative cross-sectional study	Indian (n = 27), and Caucasian (n = 27) participants	χ^2 test, unpaired and paired t-test, Bonferroni correction for multiple comparisons, Mann–Whitney U test.	- Small sample size (n = 27), potentially limiting representativeness of the diverse Indian population. - Absence of manual editing and quantification of registration accuracy. - Differences in MRI acquisition parameters between Indian scans (1.5T) and the IXI dataset (Caucasian participants), potentially introducing measurement variability. - Use of 1.5T MRI scans for compatibility with existing Western templates, despite the higher resolution of 3T imaging. - No collection of ancestral background data.	- Significant differences in brain volume were observed between Indian and Caucasian participants, necessitating an Indian-specific brain template. This underscores the importance of using population-specific templates for accurate analysis over the Montreal Neurological Institute template.
8	Tang et al. (2018),[49] China	comparative cross-sectional study	Han Chinese (n = 45) and Caucasian (n = 45) participants	Surface-based morphometric analysis, voxel-based morphometry (VBM) analysis.	- Sample limited to 45 male participants per group (total n = 90), potentially restricting generalizability. - Restriction to young male participants, limiting applicability to females and older populations. - Use of structural MRI only, without additional modalities (e.g., diffusion MRI or fMRI). - Absence of behavioral or cognitive measures to correlate with structural findings, limiting interpretation of functional significance.	- Structural differences observed between Chinese and Caucasian participants. - Differences in cortical thickness, volume, and surface area.

Table 1. Summary of characteristics of the studies included in this review – cont.

Study	Authors (year), country	Study design	Population, context	Analysis	Limitations	Key findings
9	Zahodne et al. (2015),[50] USA	cross-sectional observational study	Non-Hispanic White (n = 185), African American (n = 230), and Hispanic (n = 223) participants	Descriptive statistics, group comparisons, regression analysis, and multiple-group regression analysis.	- Cross-sectional design precluding causal interpretation between brain structure and cognitive outcomes across racial/ethnic groups. - Binary measurement of infarcts (present vs absent), potentially reducing sensitivity for detecting detailed associations with cognitive performance. - Residual confounding due to longstanding differences in socioeconomic status, education, and life experiences across racial/ethnic and immigrant groups. - Recruitment from a limited geographic area (~10 square miles in Northern Manhattan), potentially restricting generalizability and complete experiential matching across groups.	- MRI–cognition associations varied across racial/ethnic groups. - Higher white matter hyperintensities (WMH) burden associated with poorer language and executive performance in African American participants. - Weaker hippocampal volume–memory associations in Hispanic participants. - Higher cognitive scores in non-Hispanic White participants. - Group differences influenced by demographic and health factors.
10	Shaked et al. (2019),[51] USA	cross-sectional observational study	African American (n = 83) and White (n = 117) participants	Regression analysis.	- Cross-sectional design precluding causal or temporal inference between sociodemographic factors and brain structure. - Limited operationalization of socioeconomic status (SES) (poverty status and education only), without inclusion of occupation, wealth, childhood SES, or longitudinal changes. - Restricted generalizability beyond African American and White participants, with potential underrepresentation of higher-SES individuals. - Lack of correction for multiple comparisons, increasing the risk of type I error. - Absence of direct lifespan stress measures despite a stress-theoretical framework. - Incomplete MRI segmentation, excluding stress-relevant regions (e.g., hypothalamus and brainstem). - No separate analysis of key prefrontal subregions (e.g., dorsomedial vs ventromedial Prefrontal cortex (PFC)) or the anterior cingulate cortex (ACC). -ROI volumes not adjusted for total brain volume, potentially affecting volumetric comparisons.	- An interaction between SES and race was associated with differences in brain volume; high-SES White participants exhibited larger cortical volumes than African American participants and low-SES White participants. - African American participants and low-SES individuals exhibited smaller volumes in stress-related subcortical regions.
11	Waldstein et al. (2017),[52] USA	cross-sectional observational study	African American (n = 63) and White (n = 84) participants	Multiple regression analysis, t-tests, sensitivity analyses, and exploratory analyses.	- Restriction to African American and White participants, limiting generalizability to other racial/ethnic populations. - Potential geographic specificity to an urban Baltimore cohort, limiting applicability to rural or other urban settings. - Limited operationalization of SES (education and poverty status only), excluding occupational status, wealth, and detailed income measures. - Sample size limiting detection of smaller group differences and multilevel interactions (e.g., age × sex). - No correction for multiple comparisons due to exploratory design, increasing the risk of type I error.	- Low-SES African American participants exhibited a higher white matter lesion burden than low-SES White participants. - Higher-SES White participants had greater brain volumes than lower-SES White and African American participants; no SES differences were observed among African American participants. - The interaction between socioeconomic status and race suggested a complex relationship influencing brain health outcomes.

Table 1. Summary of characteristics of the studies included in this review – cont.

Study	Authors (year), country	Study design	Population, context	Analysis	Limitations	Key findings
12	Fleischman et al. (2022),[53] USA	observational study	Black participants (n = 376)	Linear-mixed effects models for cognitive change, voxel-wise linear regression for associations between deformations and cognition.	- Adjustment for cardiometabolic risk factors (e.g., hypertension, diabetes, smoking) without examination of race-relevant sociocultural moderators (e.g., educational quality, life-course experiences) influencing brain–cognition associations. - Predominantly female, highly educated sample from the Chicagoland area, limiting generalizability to the broader U.S. population of older Black adults. - Intrinsic correlations between global cognition and domain-specific abilities, limiting construct separation despite multiple assessments. - Brain morphometry assessed only at baseline, with longitudinal cognitive follow-up but no evaluation of concurrent structural and cognitive change. - Use of deformation-based morphometry (DBM) only, without complementary imaging modalities (e.g., diffusion MRI, quantitative susceptibility mapping (QSM), white matter hyperintensity burden, resting-state fMRI).	- Reduced brain tissue volume and increased cerebrospinal fluid (CSF) with age were associated with declines in global cognition, memory, and perceptual speed. - Specific brain regions, including the entorhinal cortex and hippocampi, were associated with cognitive decline. - Temporal lobe volumes were associated with episodic memory performance, whereas parietal lobe volumes were associated with perceptual speed.
13	Moonen et al. (2021),[54] USA	observational study	Black (n = 183) and White (n = 295) participants	ANOVA χ^2 tests, paired sample t-tests, multiple regression analysis.	- Selection bias due to inclusion of participants completing both MRI scans and 25- and 30-year CARDIA follow-ups; non-returning participants had poorer cardiovascular health (Life's Simple 7 (CVH-LS7) scores). - Reduced sample size due to attrition, limiting statistical power for subgroup and sex × race interaction analyses. - Cardiovascular risk factors assessed at baseline only, without longitudinal tracking. - Use of a composite CVH-LS7 score, potentially obscuring the impact of individual cardiovascular health metrics. - Limited imaging frequency (2 MRI scans over 5 years), restricting detailed characterization of brain aging trajectories. - Focus on macrostructural MRI measures without inclusion of more sensitive biomarkers of early brain pathology.	- Race- and sex-related differences were observed in baseline brain health measures. - Men exhibited greater declines in total brain and gray matter volumes compared with women. - Black participants exhibited significant differences in gray matter volume and cerebral blood flow compared with White participants.
14	Choi et al. (2024),[55] South Korea	comparative observational study	Korean (n = 1,629) and Caucasian (n = 786) participants	Regression analysis and bootstrapping.	- Scanner heterogeneity across sites for the Caucasian sample, whereas the Korean sample was acquired using a single scanner, potentially introducing measurement error and reducing statistical precision in group comparisons. - Cross-sectional study design, limiting assessment of individual brain aging trajectories. - Scanner-related variability potentially attenuating observed ethnic differences in brain aging rates, particularly within the Caucasian sample.	- Ethnicity was associated with differences in brain volume, with Korean participants generally exhibiting larger cortical structures. - Sex-related differences were observed, with certain brain structures differing in size between men and women. - A steeper aging slope was observed in Caucasian women, particularly in cortical regions. - apolipoprotein E (APOE) ε4 carrier status was associated with differential patterns of brain aging across ethnic groups.

Table 2. Joanna Briggs Institute (JBI) critical appraisal checklist for cross-sectional studies

Article (authors, year)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	JBI score (yes/total)	Overall appraisal (%)
Atilano-Barbosa and Barrios,[28] 2023	1	1	0	1	1	1	1	1	7/8	87.5
Choi et al.,[29] 2021	1	1	0	1	1	1	1	1	7/8	87.5
Kang et al.,[30] 2020	1	1	0	1	1	1	1	1	7/8	87.5
Lou et al.,[45] 2020	1	1	0	1	1	1	1	1	7/8	87.5
Lu et al.,[46] 2016	1	1	0	1	1	1	1	1	7/8	87.5
Lu et al.,[47] 2017	1	1	0	1	1	1	1	1	7/8	87.5
Rao et al.,[48] 2017	1	1	0	1	1	1	1	1	7/8	87.5
Tang et al.,[49] 2018	1	1	0	1	1	1	1	1	7/8	87.5
Shaked et al.,[51] 2019	1	1	0	1	1	1	1	1	7/8	87.5
Waldstein et al.,[52] 2017	1	1	0	1	1	1	1	1	7/8	87.5

Table 3. Joanna Briggs Institute (JBI) critical appraisal checklist for cohort studies

Article (authors, year)	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Overall score
Zahodne et al.,[50] 2015	1	1	1	1	1	1	1	1	1	1	1	11/11 (100%)
Moonen et al.,[54] 2022	1	1	1	1	1	1	1	1	1	1	1	11/11 (100%)

confounding variables. To complement Table 2[28–30,45–49,51,52] and Table 3[50,54], a domain-based summary of the potential risk of bias across the included studies is presented, based on the JBI critical appraisal tool for cross-sectional and cohort studies. Each study was evaluated across key domains, including clarity of inclusion criteria, validity of outcome measurement, and management of confounders. This visual overview provides a structured understanding of the distribution of bias risk across the evidence base.

Data analysis

Thematic analysis was performed following Braun and Clarke's six-step framework.[56] This process included familiarization with the findings, coding, and synthesis of themes related to racial differences in brain structure, diagnostic implications, and challenges associated with AI-based models. Given the high methodological heterogeneity, a meta-analysis was not feasible.[57] Instead, a qualitative synthesis was conducted to explore nuanced patterns in population-specific volumetric variability.

Results

Search outcome

Initially, 1,013 research articles, including publications in English and other languages, were identified across 4 databases (EBSCOhost, Elsevier, Clarivate, and PubMed). During the identification stage, 462 duplicate records were removed, and an additional 40 articles were excluded due to lack of full-text availability. The remaining 511 articles were screened based on titles and abstracts, and 463 were excluded as they were not relevant to the review topic. The remaining 48 articles underwent full-text review, and 31 were excluded because they did not focus on racial differences in brain volumetric outcomes. Thus, the remaining 17 relevant articles were assessed against the exclusion criteria, and 3 were excluded because the primary study population consisted of adolescents. Consequently, this scoping review included 14 articles with comparative and cross-sectional observational study designs that provided evidence relevant to the research questions. The shortlisted studies were conducted in the USA, Mexico, South Korea, China, and India (Fig. 2, Table 1).

Thematic analysis of the shortlisted studies

This scoping review synthesizes findings from 14 multinational studies published between 2014 and 2024, encompassing diverse populations, including Caucasian, East Asian, South Asian, Latinx, and African American cohorts. Specifically, findings from the shortlisted studies were categorized into 3 major themes: 1) race/ethnicity-specific differences in brain morphometry; 2) the impact of SES and race/ethnicity on brain health; and 3) the role of racial identity in cognitive aging and brain health trajectories. These themes collectively emphasize the interplay between genetic, environmental, and socioeconomic factors in shaping brain aging and disease susceptibility.

Theme 1: Race/ethnicity-specific differences in brain morphometry

Studies consistently reported structural differences across ethnic groups in cortical thickness, surface area, and subcortical volumes. The most frequently cited regions

Fig. 2. Global distribution of studies included in the review by country of origin

included the frontal, temporal, parietal, and occipital lobes, as well as subcortical structures such as the hippocampus and choroid plexus.[28–30,45–50] For instance, Chinese participants generally exhibited larger temporal and cingulate cortices but smaller frontal and parietal regions compared with Caucasian participants.[49] Similarly, significant morphometric differences were observed between Chinese and Indian groups, particularly in gray matter distribution within the frontal, temporal, and occipital lobes.[47] Studies comparing East Asian groups (Koreans and Uyghurs) with White participants reported widespread differences in cortical thickness and surface area, especially in language-associated regions such as the frontoparietal operculum.[29,30,46] In U.S.-based samples, African American participants exhibited distinct patterns of white matter and subcortical volume compared with White participants,[28] with larger hippocampal volumes associated with better memory outcomes among non-Hispanic White individuals.[50] These group-level differences in brain volume should be understood as arising from interacting factors such as ancestry, environmental exposures, education, and systemic inequalities, rather than as evidence of inherent biological distinctions. This perspective supports the view that population-level differences are shaped by environmental influences over time rather than by innate biological traits.

Theme 2: Impact of socioeconomic status and race/ethnicity on brain health

The interplay between race/ethnicity, SES, and brain health is a significant factor in understanding neurodevelopment and cognitive aging. Disparities in education, income, access to healthcare, and employment stability influence brain structure and function, with cumulative exposure to stressors exacerbating racial disparities in neurological health (Fig. 3).[51,52] Waldstein et al.[52] found that higher SES was associated with neuroprotection against cognitive decline and cerebrovascular damage, predominantly among White populations, whereas African American participants derived fewer protective benefits from SES advantages. This finding suggests that economic gains alone may not counteract the effects of chronic stressors that disproportionately affect racial minority groups. Zahodne et al.[50] identified racial disparities in WMH, an MRI marker of cerebrovascular health. African American participants exhibited larger WMH volumes, which were strongly associated with poorer executive function, whereas this pattern was not observed in non-Hispanic White participants.[50] Increased WMH burden was associated with lower psychomotor speed and poorer executive function, reflecting overall cognitive status and brain health.[58] Fleischman et al.[53] found that lower hippocampal volume was associated with episodic memory deficits in aging Black adults. Despite these well-documented disparities, current AI-driven diagnostic models in neuroimaging often fail to incorporate SES-related variables, potentially exacerbating healthcare inequities.

Theme 3: The impact of racial identity on cognitive aging and brain health trajectories

Racial identity influences not only baseline brain structure but also the trajectory of cortical atrophy during aging. Studies suggest that cortical thinning accelerates with age, typically worsening after 70 years of age,[59] although these patterns vary significantly across racial and gender groups.

Fig. 3. Heatmap illustrating the interaction between socioeconomic status (SES), race, and brain health outcomes

High SES Whites: SES has a low impact on stroke risk, cognitive decline, and white matter lesions.
Low SES Whites: Moderate impact observed in all cognitive outcomes.
High SES African Americans: Moderate to high impact, indicating the importance of SES in this group.
Low SES African Americans: Very high impact on stroke risk, cognitive decline, and white matter lesions, underscoring severe effects of low SES.

Choi et al.[29] observed that Caucasian women had steeper rates of cortical atrophy compared with Korean women, suggesting the influence of potential cultural or environmental protective factors. Moonen et al.[54] reported faster gray matter decline in Black men than in White individuals or Black women, likely reflecting chronic stress exposure and systemic inequalities. These findings are supported by studies indicating that brain regions involved in language and cognition, such as the frontal and temporal lobes, are particularly susceptible to social determinants of health.[60,61]

Discussion

This scoping review examines the need for population-sensitive neuroimaging standards by synthesizing evidence on differences in brain morphology and their implications for AI-based diagnostics, while critically addressing the concept of biological essentialism.

Influence of biological and SES factors and associated risk of bias

While biological distinctions in brain morphology may support the development of population-sensitive neuroimaging reference standards, these differences are shaped by evolutionary adaptations[30,62] and broader factors, such as SES and systemic discrimination, which accelerate brain aging in marginalized groups, as evidenced by accelerated gray matter decline among Black adults.[34,54] This suggests that the observed differences in brain morphology are more accurately attributed to socioeconomic and environmental determinants than to innate biological variation.[34,63] Similarly, African American participants demonstrate a stronger association between larger WMH volumes and poorer cognitive function compared with non-Hispanic White individuals.[50] Documented evidence of ethnic disparities in biomarkers of AD and PD highlights the risk of misclassification if AI models are trained on homogeneous datasets, potentially leading to inaccurate diagnoses in diverse populations.[29,50]

Advantages and limitations of population-sensitive templates

Developing population-sensitive MRI templates and normative brain volumes (NBVs) offers potential advantages for diagnostic accuracy. Studies by Choi et al.[29,55] demonstrated that population-specific templates significantly enhance diagnostic precision and reduce prediction errors in neurodegenerative disease models. For example, an MRI template developed for Indian individuals was tested on MRI scans from an Indian population, yielding better alignment compared with the standard MRI template based on a Caucasian population.[48]

In addition, Caucasian normative brain volume references were unable to accurately classify Korean patients with AD, as AD-vulnerable regions are larger in cognitively normal older Korean adults.[29]

Subsequent studies found that ethnicity-adjusted norms for brain volumes improved the diagnostic accuracy of AD.[55] Nathoo et al.[64] similarly highlighted substantial disparities in multiple sclerosis lesion burden and clinical outcomes, with poorer outcomes observed among African American and Latin American populations compared with non-Hispanic White individuals. The same study also reported increased lesion load and worse clinical outcomes among Latin American patients, further emphasizing health disparities across underrepresented populations. In contrast, Hedderich et al.[65] found that normative brain volumes have a limited impact on diagnostic accuracy in neurodegenerative diseases. These variations

underscore the need for population-sensitive neuroimaging models that account for biological and socioeconomic influences to improve diagnostic precision.

Integrating race/ethnicity and SES in AI models and their ethical considerations

Although integrating race and SES into AI models offers potential advantages, addressing issues of transparency and ethical governance is essential. Reinforcing racial essentialism and embedding bias in clinical AI systems – particularly through misinterpreting race as a biological determinant – may result in the reinforcement of racial stereotypes, breaches of privacy, and the exacerbation of healthcare inequities. It is crucial to distinguish biologically based variability from disparities arising from systemic inequities, such as structural racism, limited access to healthcare, and environmental exposures, in order to prevent the reinforcement of biological essentialism. Therefore, population-sensitive standards must be developed alongside appropriate safeguards, including the incorporation of socioeconomic variables, robust informed consent procedures, and continuous bias monitoring, to ensure equitable implementation in clinical practice. Given the critical role of SES in brain health, its integration as a core variable in neuroimaging datasets – using tools such as the HOUSES index, which measures individual socioeconomic status based on housing characteristics such as cost and crowding – is essential.[66]

To mitigate these risks, researchers and clinicians must ensure transparency in data collection methodologies, obtain explicit informed consent from participants, and implement robust ethical oversight mechanisms. Providing clear information about the populations included and their demographic characteristics can help clinicians make better-informed decisions when using AI-based neuroimaging tools. Artificial intelligence-based diagnostic tools should serve as supportive aids rather than substitutes for clinical judgment, ensuring that radiologists and neurologists retain primary responsibility for interpreting neuroimaging results within the broader clinical context (Fig. 4).

Artificial intelligence-based neuroimaging platforms should incorporate a structured clinician feedback loop to ensure continuous improvement and ethical oversight. Clinicians should be able to flag suspected diagnostic discrepancies through a secure, anonymized interface. Expert reviewers should then validate these flagged cases, and confirmed discrepancies should inform model retraining pipelines. This iterative process would enhance model calibration, facilitate bias detection, and promote accountability in the deployment of clinical AI systems. Developers can implement broader ethical safeguards by conducting subgroup-specific performance audits, applying bias quantification metrics, and validating models using diverse population datasets. Explainable AI (XAI) techniques, such as SHAP (Shapley Additive Explanations) and LIME (Local Interpretable Model-Agnostic Explanations), can help clarify model decision-making processes, thereby reducing the "black box" effect and highlighting the influence of input features such as SES and race/ethnicity. Consent frameworks should clearly define data provenance, usage parameters, and the implications of algorithmic outputs to ensure transparency and participant autonomy. Finally, clinicians must receive targeted training to interpret AI predictions in the context of population metadata and social determinants of health, reinforcing their essential role in context-aware, multidisciplinary diagnostic workflows.

Limitations of the study

This scoping review, by design, does not include a critical appraisal of bias or a meta-analysis, which would provide the statistical precision and pooled effect estimates typically reported in systematic reviews. Nevertheless, adherence to the PRISMA-ScR framework mitigated selection bias by ensuring a structured approach to study identification, screening, and inclusion.[42] As noted in the "Study heterogeneity" section, methodological inconsistencies in MRI protocols, software, and adjustments for confounding variables pose challenges to generalizability. Furthermore, a substantial proportion of the included studies employed cross-sectional designs, limiting the ability to establish causal relationships or assess longitudinal changes. Variations in sample size (ranging from 30 to 2,537 participants) also constrain generalizability. The present review primarily emphasized structural and socio-structural dimensions without extensively addressing genetic predispositions or environmental exposures, including air pollution, nutrition, and early-life stress.

Conclusions

This scoping review synthesizes evidence of population-sensitive variability in regional brain morphology across racial, ethnic, and socioeconomic groups in both health and disease. These differences have important implications for the interpretation of neuroimaging findings and the development of AI-based diagnostic models. The review highlights persistent gaps in demographic diversity within current neuroimaging datasets, particularly in commercial and research-based AI tools. Addressing these gaps requires studies that incorporate larger, socio-demographically diverse populations and account for structural determinants of health. Developing population-sensitive neuroimaging templates is critical for improving diagnostic accuracy and advancing healthcare equity; however, this effort must also consider practical challenges, including data harmonization, potential misclassification, and ethical risks such as the inadvertent reinforcement of racial bias. To navigate these complexities, coordinated efforts

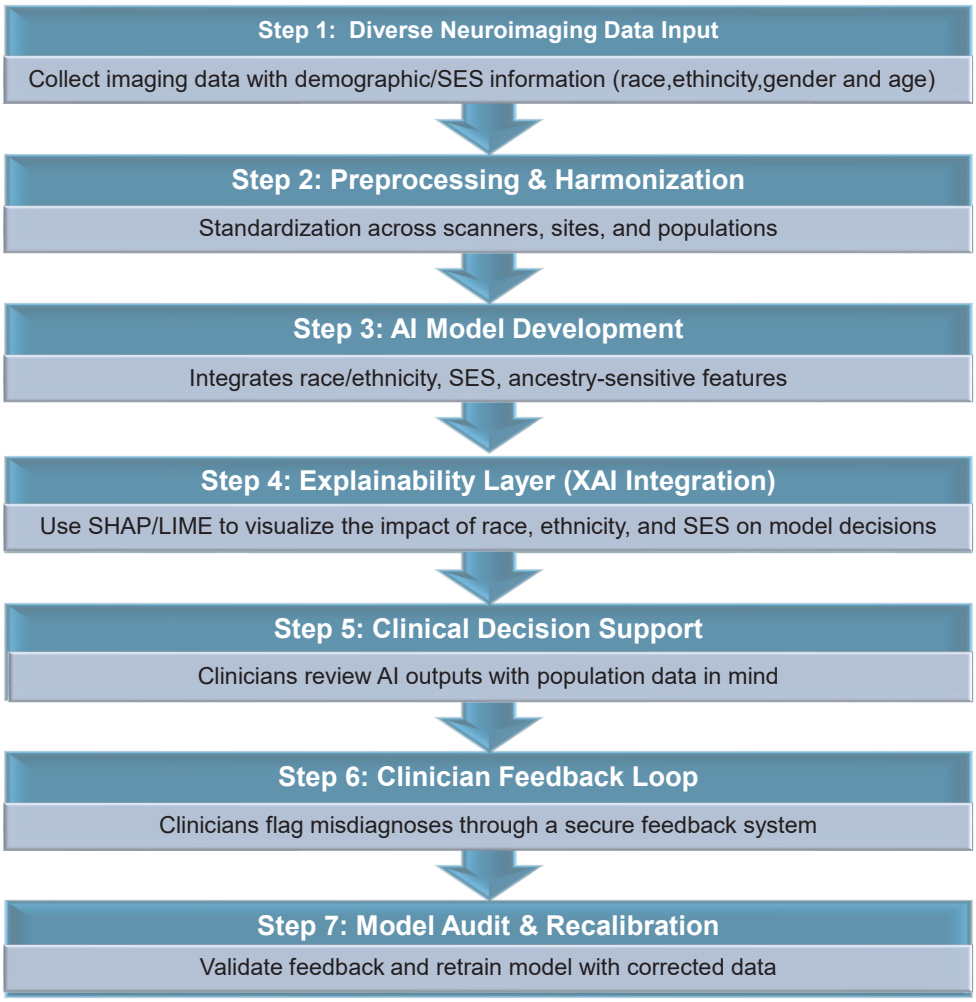

Fig. 4. Future implications of the review findings

SES – socioeconomic status; AI – artifical intelligence.

at national and international levels are necessary, alongside the rigorous implementation of ethical safeguards. These should include transparency in data provenance, the inclusion of socioeconomic variables, and interpretability mechanisms within AI tools. Explainable AI techniques can be incorporated to visualize how input features – including race, ethnicity, and socioeconomic indicators – contribute to diagnostic predictions. These methods help quantify feature attribution, promote transparency in model behavior, and enable clinicians to assess whether outputs are influenced by spurious or biased correlations. The proposed implementation approaches are consistent with emerging recommendations to integrate structural and functional neuroimaging features into personalized, context-aware diagnostic frameworks.[7,65] Population-sensitive standards and inclusive models may help reduce bias, misclassification, and inequities in AI-driven neuroimaging. Interpretation of intergroup brain differences must remain grounded in a socio-structural framework to prevent the misrepresentation of race as a biological essence. Crucially, group-level morphological differences should be interpreted within the broader context of social determinants of health rather than as biologically inherent traits, in order to avoid reinforcing essentialist narratives.

Recommendations

To address the clinical, ethical, and research challenges identified in this review, the following strategies are recommended:

– Fostering international partnerships and data-sharing initiatives to create large, representative neuroimaging datasets that capture population diversity. These datasets could be stratified by race/ethnicity, age, and sex to improve diagnostic precision and minimize bias in AI applications.

– Incorporating validated socioeconomic indicators (e.g., HOUSES index, education, housing quality) and ensuring subgroup performance auditing during model development.

– Employing explainability (XAI) methods, such as SHAP and LIME, to enhance transparency in algorithmic decision-making.

– Integrating population-sensitive neuroanatomical features (e.g., ethnicity-adjusted volumetrics, SES-modulated patterns) into AI models while contextualizing them as proxies for lived structural and social experiences rather than fixed biological traits. This approach safeguards against the misinterpretation of race as a biological essence

and aligns diagnostic models with equity-focused precision medicine.

– Employing federated learning approaches to preserve data privacy while improving inclusivity across geographic and demographic populations.[39]

– Implementing ethical AI toolkits to detect and address algorithmic bias at all stages of model deployment.

– Establishing a structured feedback mechanism through which clinicians can report AI-generated diagnostic discrepancies, allowing for systematic monitoring, auditing, and model retraining. This mechanism should include a secure, anonymized reporting interface and a validation pipeline for incorporating clinical corrections into future model updates.

– Developing future models that include metadata fields capturing key social determinants of health.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Srinivasa Rao Bolla https://orcid.org/0000-0002-2644-5169
Joseph Uy Almazan https://orcid.org/0000-0001-5148-6889
Rauan Satbekov https://orcid.org/0000-0007-4817-6952
Syed Hani Abidi https://orcid.org/0000-0001-9497-0902
Dinara Jumadilova https://orcid.org/0000-0001-7958-7956
Kamila Mussabekova https://orcid.org/0009-0008-0201-4496
Surendra Maharjan https://orcid.org/0000-0002-0572-1611

References

1. Nichols E, Steinmetz JD, Vollset SE, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. *Lancet Public Health*. 2022;7(2):e105–e125. doi:10.1016/S2468-2667(21)00249-8
2. Bajaj S, Khunte M, Moily NS, et al. Value proposition of FDA-approved artificial intelligence algorithms for neuroimaging. *J Am Coll Radiol*. 2023;20(12):1241–1249. doi:10.1016/j.jacr.2023.06.034
3. Razdan P, Das A, Habiba S, Doley S, Tiwari DA, Hazari P. Knowledge, perception and attitude of dentists regarding the role of artificial intelligence in the field of pediatric dentistry: An online questionnaire study [published online as ahead of print on February 19, 2025]. *Dent Med Probl*. 2025. doi:10.17219/dmp/183901
4. Szczepanowski R, Uchmanowicz I, Pasieczna AH, et al. Application of machine learning in predicting frailty syndrome in patients with heart failure. *Adv Clin Exp Med*. 2024;33(3):309–315. doi:10.17219/acem/184040
5. Uysal HA, Poyraz T, Gulluoglu H, Idiman F, Idiman E. An artificial intelligence model for Lhermitte's sign in patients with pediatric-onset multiple sclerosis: A follow-up study. *Adv Clin Exp Med*. 2025;34(2):165–177. doi:10.17219/acem/196466
6. Armstrong NM, An Y, Shin JJ, et al. Associations between cognitive and brain volume changes in cognitively normal older adults. *NeuroImage*. 2020;223:117289. doi:10.1016/j.neuroimage.2020.117289
7. Alkan E, Evans SLH. Relationships between cognitive performance, clinical insight and regional brain volumes in schizophrenia. *Schizophrenia*. 2022;8(1):33. doi:10.1038/s41537-022-00243-x
8. Matías-Guiu JA, Cortés-Martínez A, Montero P, et al. Identification of cortical and subcortical correlates of cognitive performance in multiple sclerosis using voxel-based morphometry. *Front Neurol*. 2018;9:920. doi:10.3389/fneur.2018.00920
9. Koshiyama D, Fukunaga M, Okada N, et al. Role of subcortical structures on cognitive and social function in schizophrenia. *Sci Rep*. 2018;8(1):1183. doi:10.1038/s41598-017-18950-2
10. Battaglia S, Schmidt A, Hassel S, Tanaka M. Editorial: Case reports in neuroimaging and stimulation. *Front Psychiatry*. 2023;14:1264669. doi:10.3389/fpsyt.2023.1264669
11. Battaglia S, Nazzi C, Di Fazio C, Borgomaneri S. The role of pre-supplementary motor cortex in action control with emotional stimuli: A repetitive transcranial magnetic stimulation study. *Ann N Y Acad Sci*. 2024;1536(1):151–166. doi:10.1111/nyas.15145
12. Hartikainen P, Räsänen J, Julkunen V, et al. Cortical thickness in frontotemporal dementia, mild cognitive impairment, and Alzheimer's disease. *J Alzheimers Dis*. 2012;30(4):857–874. doi:10.3233/JAD-2012-112060
13. Lerch JP, Pruessner JC, Zijdenbos A, Hampel H, Teipel SJ, Evans AC. Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy. *Cereb Cortex*. 2005;15(7):995–1001. doi:10.1093/cercor/bhh200
14. Querbes O, Aubry F, Pariente J, et al. Early diagnosis of Alzheimer's disease using cortical thickness: Impact of cognitive reserve. *Brain*. 2009;132(8):2036–2047. doi:10.1093/brain/awp105
15. Zarei M, Ibarretxe-Bilbao N, Compta Y, et al. Cortical thinning is associated with disease stages and dementia in Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2013;84(8):875–882. doi:10.1136/jnnp-2012-304126
16. Stoebner ZA, Hett K, Lyu I, et al. Comprehensive shape analysis of the cortex in Huntington's disease. *Hum Brain Mapp*. 2023;44(4):1417–1431. doi:10.1002/hbm.26125
17. Walhout R, Westeneng HJ, Verstraete E, et al. Cortical thickness in ALS: Towards a marker for upper motor neuron involvement. *J Neurol Neurosurg Psychiatry*. 2015;86(3):288–294. doi:10.1136/jnnp-2013-306839
18. Li Q, Zhao Y, Chen Z, et al. Meta-analysis of cortical thickness abnormalities in medication-free patients with major depressive disorder. *Neuropsychopharmacology*. 2020;45(4):703–712. doi:10.1038/s41386-019-0563-9
19. Suh JS, Minuzzi L, Raamana PR, et al. An investigation of cortical thickness and antidepressant response in major depressive disorder: A CAN-BIND study report. *Neuroimage Clin*. 2020;25:102178. doi:10.1016/j.nicl.2020.102178
20. Zhao Y, Zhang Q, Shah C, et al. Cortical thickness abnormalities at different stages of the illness course in schizophrenia: A systematic review and meta-analysis. *JAMA Psychiatry*. 2022;79(6):560. doi:10.1001/jamapsychiatry.2022.0799
21. Hanford LC, Nazarov A, Hall GB, Sassi RB. Cortical thickness in bipolar disorder: A systematic review. *Bipolar Disord*. 2016;18(1):4–18. doi:10.1111/bdi.12362
22. Niu M, Wang Y, Jia Y, et al. Common and specific abnormalities in cortical thickness in patients with major depressive and bipolar disorders. *EBioMedicine*. 2017;16:162–171. doi:10.1016/j.ebiom.2017.01.010
23. Shukla A, Tiwari R, Tiwari S. Analyzing subcortical structures in Alzheimer's disease using ensemble learning. *Biomed Signal Process Control*. 2024;87:105407. doi:10.1016/j.bspc.2023.105407
24. García-Marín LM, Reyes-Pérez P, Diaz-Torres S, et al. Shared molecular genetic factors influence subcortical brain morphometry and Parkinson's disease risk. *NPJ Parkinsons Dis*. 2023;9(1):73. doi:10.1038/s41531-023-00515-y
25. Okada N, Fukunaga M, Miura K, et al. Subcortical volumetric alterations in four major psychiatric disorders: A mega-analysis study of 5604 subjects and a volumetric data-driven approach for classification. *Mol Psychiatry*. 2023;28(12):5206–5216. doi:10.1038/s41380-023-02141-9
26. Shi J, Guo H, Liu S, et al. Subcortical brain volumes relate to neurocognition in first-episode schizophrenia, bipolar disorder, major depression disorder, and healthy controls. *Front Psychiatry*. 2022;12:747386. doi:10.3389/fpsyt.2021.747386
27. Tu PC, Chang WC, Chen MH, et al. Identifying common and distinct subcortical volumetric abnormalities in 3 major psychiatric disorders: A single-site analysis of 640 participants. *J Psychiatry Neurosci*. 2022;47(3):E230–E238. doi:10.1503/jpn.210154
28. Atilano-Barbosa D, Barrios FA. Brain morphological variability between whites and African Americans: The importance of racial identity in brain imaging research. *Front Integr Neurosci*. 2023;17:1027382. doi:10.3389/fnint.2023.1027382

29. Choi YY, Lee JJ, Choi KY, et al. Multi-racial normative data for lobar and subcortical brain volumes in old age: Korean and Caucasian norms may be incompatible with each other. *Front Aging Neurosci*. 2021;13:675016. doi:10.3389/fnagi.2021.675016
30. Kang DW, Wang SM, Na HR, et al. Differences in cortical structure between cognitively normal East Asian and Caucasian older adults: A surface-based morphometry study. *Sci Rep*. 2020;10(1):20905. doi:10.1038/s41598-020-77848-8
31. Austin TR, Nasrallah IM, Erus G, et al. Association of brain volumes and white matter injury with race, ethnicity, and cardiovascular risk factors: The multi-ethnic study of atherosclerosis. *J Am Heart Assoc*. 2022;11(7):e023159. doi:10.1161/JAHA.121.023159
32. Peper JS, Brouwer RM, Boomsma DI, Kahn RS, Hulshoff Pol HE. Genetic influences on human brain structure: A review of brain imaging studies in twins. *Hum Brain Mapp*. 2007;28(6):464–473. doi:10.1002/hbm.20398
33. Geronimus AT, Pearson JA, Linnenbringer E, et al. Race-ethnicity, poverty, urban stressors, and telomere length in a Detroit community-based sample. *J Health Soc Behav*. 2015;56(2):199–224. doi:10.1177/0022146515582100
34. Turney IC, Lao PJ, Rentería MA, et al. Brain aging among racially and ethnically diverse middle-aged and older adults. *JAMA Neurol*. 2023;80(1):73. doi:10.1001/jamaneurol.2022.3919
35. Hussain MA, Grant PE, Ou Y. Inferring neurocognition using artificial intelligence on brain MRIs. *Front Neuroimaging*. 2024;3:1455436. doi:10.3389/fnimg.2024.1455436
36. Koçak B, Ponsiglione A, Stanzione A, et al. Bias in artificial intelligence for medical imaging: Fundamentals, detection, avoidance, mitigation, challenges, ethics, and prospects. *Diagn Interv Radiol*. 2024;31(2):75–88. doi:10.4274/dir.2024.242854
37. Van Leeuwen KG, Schalekamp S, Rutten MJCM, Van Ginneken B, De Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. *Eur Radiol*. 2021;31(6):3797–3804. doi:10.1007/s00330-021-07892-z
38. Kocak B, Baessler B, Cuocolo R, Mercaldo N, Pinto Dos Santos D. Trends and statistics of artificial intelligence and radiomics research in radiology, nuclear medicine, and medical imaging: Bibliometric analysis. *Eur Radiol*. 2023;33(11):7542–7555. doi:10.1007/s00330-023-09772-0
39. Ricard JA, Parker TC, Dhamala E, Kwasa J, Allsop A, Holmes AJ. Confronting racially exclusionary practices in the acquisition and analyses of neuroimaging data. *Nat Neurosci*. 2023;26(1):4–11. doi:10.1038/s41593-022-01218-y
40. Zhang M, Wang Y, Lv M, et al. Trends and hotspots in global radiomics research: A bibliometric analysis. *Technol Cancer Res Treat*. 2024;23:15330338241235769. doi:10.1177/15330338241235769
41. Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. *Int J Soc Res Methodol*. 2005;8(1):19–32. doi:10.1080/1364557032000119616
42. Levac D, Colquhoun H, O'Brien KK. Scoping studies: Advancing the methodology. *Implement Sci*. 2010;5(1):69. doi:10.1186/1748-5908-5-69
43. Munn Z, Peters MDJ, Stern C, Tufanaru C, McArthur A, Aromataris E. Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med Res Methodol*. 2018;18(1):143. doi:10.1186/s12874-018-0611-x
44. Moola S, Munn Z, Tufunaru C, et al. Systematic reviews of aetiology and risk. In: Aromataris E, Lockwood C, Porritt K, Pilla B, Jordan Z, eds. *JBI Manual for Evidence Synthesis*. Adelaide, Australia: Joanna Briggs Institute (JBI); 2024. doi:10.46658/JBIMES-24-06
45. Lou Y, Zhao L, Yu S, et al. Brain asymmetry differences between Chinese and Caucasian populations: A surface-based morphometric comparison study. *Brain Imaging Behav*. 2020;14(6):2323–2332. doi:10.1007/s11682-019-00184-7
46. Lu J, Jiang C, Wang J, Jia W. Comprehensive cortical thickness and surface area comparison between young Uyghur and Han Chinese cohorts. *Magn Reson Imaging*. 2016;34(8):1043–1049. doi:10.1016/j.mri.2016.03.018
47. Lu J, Peng B, Saxena A, et al. Examining population differences in cerebral morphometry between Chinese and Indian undergraduate students. *Neurosci Lett*. 2017;636:290–297. doi:10.1016/j.neulet.2016.11.021
48. Rao NP, Jeelani H, Achalia R, et al. Population differences in brain morphology: Need for population specific brain template. *Psychiatry Res Neuroimaging*. 2017;265:1–8. doi:10.1016/j.pscychresns.2017.03.018
49. Tang Y, Zhao L, Lou Y, et al. Brain structure differences between Chinese and Caucasian cohorts: A comprehensive morphometry study. *Hum Brain Mapp*. 2018;39(5):2147–2155. doi:10.1002/hbm.23994
50. Zahodne L, Manly J, Narkhede A, et al. Structural MRI predictors of late-life cognition differ across African Americans, Hispanics, and Whites. *Curr Alzheimer Res*. 2015;12(7):632–639. doi:10.2174/1567205012666150530203214
51. Shaked D, Millman ZB, Moody DLB, et al. Sociodemographic disparities in corticolimbic structures. *PLoS One*. 2019;14(5):e0216338. doi:10.1371/journal.pone.0216338
52. Waldstein SR, Dore GA, Davatzikos C, et al. Differential associations of socioeconomic status with global brain volumes and white matter lesions in African American and white adults: The HANDLS SCAN Study. *Psychosom Med*. 2017;79(3):327–335. doi:10.1097/PSY.0000000000000408
53. Fleischman DA, Arfanakis K, Leurgans SE, et al. Associations of deformation-based brain morphometry with cognitive level and decline within older Blacks without dementia. *Neurobiol Aging*. 2022;111:35–43. doi:10.1016/j.neurobiolaging.2021.11.003
54. Moonen JEF, Nasrallah IM, Detre JA, et al. Race, sex, and mid-life changes in brain health: Cardia MRI substudy. *Alzheimers Dement*. 2022;18(12):2428–2437. doi:10.1002/alz.12560
55. Choi YY, Lee JJ, Te Nijenhuis J, et al. Multi-ethnic norms for volumes of subcortical and lobar brain structures measured by Neuro I: Ethnicity may improve the diagnosis of Alzheimer's disease. *J Alzheimers Dis*. 2024;99(1):223–240. doi:10.3233/JAD-231182
56. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77–101. doi:10.1191/1478088706qp063oa
57. Vaismoradi M, Turunen H, Bondas T. Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nurs Health Sci*. 2013;15(3):398–405. doi:10.1111/nhs.12048
58. for the ALFA Study; Brugulat-Serrat A, Salvadó G, Sudre CH, et al. Patterns of white matter hyperintensities associated with cognition in middle-aged cognitively healthy individuals. *Brain Imaging Behav*. 2020;14(5):2012–2023. doi:10.1007/s11682-019-00151-2
59. Walhovd KB, Fjell AM, Giedd J, Dale AM, Brown TT. Through thick and thin: A need to reconcile contradictory results on trajectories in human cortical development. *Cereb Cortex*. 2016;27(2):1472–1481. doi:10.1093/cercor/bhv301
60. Eckert MA, Vaden KI, Iuricich F; Dyslexia Data Consortium. Cortical asymmetries at different spatial hierarchies relate to phonological processing ability. *PLoS Biol*. 2022;20(4):e3001591. doi:10.1371/journal.pbio.3001591
61. Novén M, Olsson H, Helms G, Horne M, Nilsson M, Roll M. Cortical and white matter correlates of language-learning aptitudes. *Hum Brain Mapp*. 2021;42(15):5037–5050. doi:10.1002/hbm.25598
62. Kweon H, Aydogan G, Dagher A, et al. Human brain anatomy reflects separable genetic and environmental components of socioeconomic status. *Sci Adv*. 2022;8(20):eabm2923. doi:10.1126/sciadv.abm2923
63. Gavett BE, Fletcher E, Harvey D, et al. Ethnoracial differences in brain structure change and cognitive change. *Neuropsychology*. 2018;32(5):529–540. doi:10.1037/neu0000452
64. Nathoo N, Zeydan B, Neyal N, Chelf C, Okuda DT, Kantarci OH. Do magnetic resonance imaging features differ between persons with multiple sclerosis of various races and ethnicities? *Front Neurol*. 2023;14:1215774. doi:10.3389/fneur.2023.1215774
65. Hedderich DM, Schmitz-Koep B, Schuberth M, et al. Impact of normative brain volume reports on the diagnosis of neurodegenerative dementia disorders in neuroradiology: A real-world, clinical practice study. *Front Aging Neurosci*. 2022;14:971863. doi:10.3389/fnagi.2022.971863
66. Juhn YJ, Ryu E, Wi CI, et al. Assessing socioeconomic bias in machine learning algorithms in health care: A case study of the HOUSES index. *J Am Med Inform Assoc*. 2022;29(7):1142–1151. doi:10.1093/jamia/ocac052

Gastrointestinal manifestations of systemic lupus erythematosus (SLE): A comprehensive literature review

Wojciech Bajurny[1,B–F], Julia Grabowska[1,B–F], Amelia Pielech[1,B–F], Natalia Struzik[1,B–F], Jakub A. Mastalerz[1,A,C,E,F], Magdalena Szmyrka[2,A,E,F]

[1] student, Faculty of Medicine, Wroclaw Medical University, Poland
[2] Department of Rheumatology, 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(5):905–918

Address for correspondence
Jakub A. Mastalerz
E-mail: jakub.mastalerz@student.umw.edu.pl

Funding sources
None declared

Conflict of interest
None declared

Received on June 16, 2025
Reviewed on July 27, 2025
Accepted on August 5, 2025

Published online on May 20, 2026

Abstract

Systemic lupus erythematosus (SLE) is a complex autoimmune disease characterized by a broad spectrum of clinical manifestations, including gastrointestinal (GI) involvement. Although the joints, skin, and kidneys are most commonly affected, GI manifestations are frequently underrecognized despite their potential to significantly influence patient outcomes. Eleven major GI manifestations have been identified and described: oral ulcers, lupus enteritis (LE), lupus peritonitis (LP), mesenteric vasculitis, mesenteric thrombosis, protein-losing enteropathy (PLE), intestinal pseudo-obstruction (IPO), lupus pancreatitis, lupus hepatitis (LH), gastroesophageal reflux disease (GERD), and medication-related adverse effects.

These manifestations may present with nonspecific symptoms, such as abdominal pain, diarrhea, vomiting, and weight loss, which often complicate timely diagnosis. Imaging modalities, particularly contrast-enhanced computed tomography, together with serological markers, including antinuclear antibodies (ANA) and complement levels, play a central role in diagnosis. Corticosteroids remain the cornerstone of treatment, whereas immunosuppressive agents and biologic therapies are reserved for refractory cases. Medication-induced GI adverse effects, particularly those associated with glucocorticoids, nonsteroidal anti-inflammatory drugs (NSAIDs), and immunosuppressive agents such as azathioprine (AZA) and cyclophosphamide (CPA), also represent important contributors to GI pathology in patients with SLE. Gastrointestinal involvement in SLE is heterogeneous and may be severe. Increased awareness, early recognition, and individualized treatment strategies are essential for improving patient outcomes. Further research is required to establish standardized diagnostic criteria and therapeutic guidelines for GI manifestations in SLE.

Key words: lupus erythematosus, rheumatology, gastroenterology

Cite as
Bajurny W, Grabowska J, Pielech A, Struzik N, Mastalerz JA, Szmyrka M. Gastrointestinal manifestations of systemic lupus erythematosus (SLE): A comprehensive literature review. *Adv Clin Exp Med*. 2026;35(5):905–918. doi:10.17219/acem/209024

DOI
10.17219/acem/209024

Highlights

- Systemic lupus erythematosus (SLE) is associated with multiple gastroenterological manifestations.
- Early gastrointestinal symptoms may support earlier diagnosis of SLE.
- Targeted imaging improves detection of mesenteric vasculitis and lupus enteritis in SLE patients.

Introduction

Systemic lupus erythematosus (SLE) is one of the most prevalent multisystem autoimmune disorders. It is characterized by the production of a broad spectrum of antinuclear antibodies (ANA) and the deposition of immune complexes in various organs, leading to inflammation and tissue damage. Although SLE most commonly affects the joints, skin, and kidneys, it may also manifest with respiratory, neurological, or gastrointestinal (GI) symptoms, reflecting the systemic nature of the disease (Fig. 1).[1,2]

The global incidence of SLE is estimated at 5.14 cases per 100,000 person-years, including 8.82 cases among women and 1.53 cases among men. The estimated worldwide prevalence is 43.7 cases per 100,000 individuals, with approx. 78.73/100,000 women and 9.26/100,000 men affected. Most documented cases have been reported in more developed regions, particularly in Europe and North America. However, these data should be interpreted with caution due to the lack of reliable epidemiological data from certain countries.[3] It is also important to note that SLE is most commonly diagnosed in young adults. Moreover, earlier disease onset is associated with a more aggressive clinical course. This observation highlights the importance of early and accurate diagnosis of the disease.[4]

As mentioned previously, SLE may also involve the GI tract among the many organ systems affected by the disease. Abdominal pain and other GI symptoms are associated with a broad differential diagnosis.[5] Because nonspecific GI manifestations may represent the initial clinical presentation of SLE, and delayed diagnosis may lead to a more severe disease course, accurate classification and comprehensive understanding of the epidemiology, clinical manifestations, diagnostic approaches, and treatment of GI involvement in SLE are essential.[6] This paper presents a scoping review of the gastroenterological manifestations of SLE. We systematically mapped the epidemiology, pathophysiology, clinical features, diagnostic approaches, and management strategies related to eleven distinct GI manifestations: oral ulcers; gastroesophageal reflux disease (GERD); enteritis; peritonitis; vasculitis; thrombosis; protein-losing enteropathy (PLE); intestinal pseudo-obstruction (IPO); pancreatitis; hepatitis; and drug-related adverse effects (Table 1).

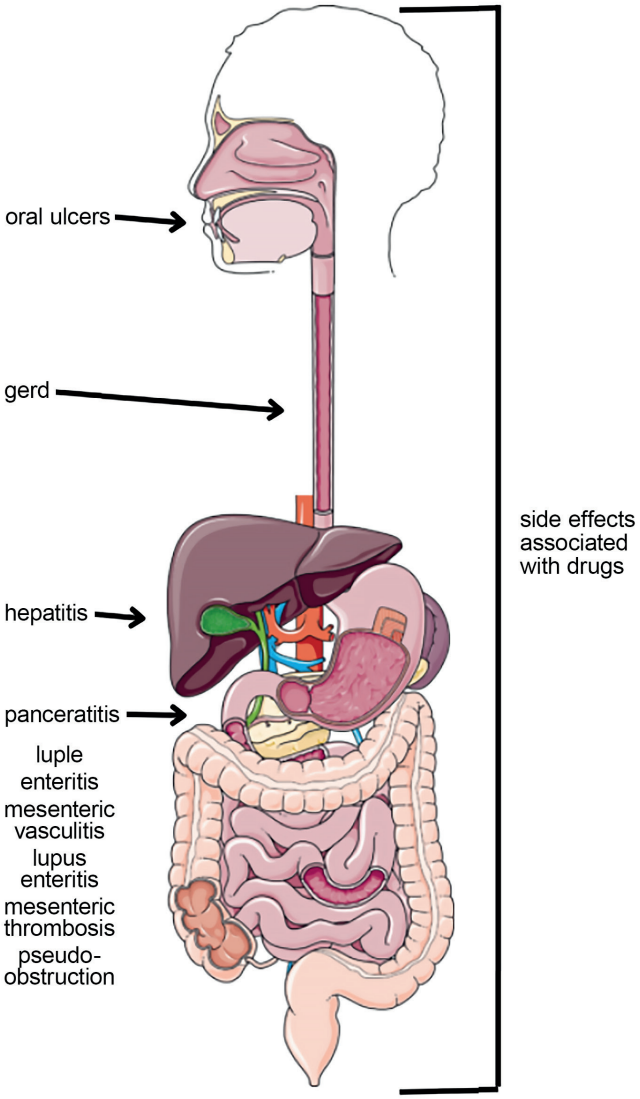

Fig. 1. Lupus may manifest throughout the entire gastrointestinal tract

Objectives

This scoping review aims to systematically map current evidence on the epidemiology, pathophysiology, clinical presentation, diagnosis, and management of gastrointestinal manifestations in SLE. By comprehensively describing eleven distinct GI manifestations, we seek to raise clinical awareness, highlight gaps in standardized diagnostic and therapeutic guidelines, and identify priorities for future research.

Table 1. Summary of the most common gastroenterological manifestations of SLE

Manifestation of SLE	Prevalence of each manifestation	Common symptoms	Diagnostic markers	Imaging findings	Standard treatments
GERD	almost half of patients with SLE	regurgitation (27.3%), heartburn (24.6%)	inflammatory response in the esophageal muscles	ischemic vasculitis, muscle atrophy	proton pump inhibitors, prokinetic agents, and high-dose H_2-blockers
PLE	rare	susceptibility to infections, peripheral edema, ascites, abdominal pain	hypoproteinemia, hypoalbuminemia, lymphopenia	^{99m}Tc-HSA demonstrating protein leakage into the bowel lumen	glucocorticoids and immunosuppressants
Mesenteric vasculitis	0.2–9.7%	abdominal pain, tenderness, rectal bleeding, nausea, vomiting, diarrhea	low complement (C3, C4) levels, CRP	bowel wall thickening, bowel loop dilation, "target sign" on CT	anti-inflammatory and immunosuppressive therapy
Mesenteric vascular thrombosis	very rare	abdominal pain, vomiting	elevated white blood cell count, D-dimer, CRP, and lactic acid levels	bowel dilatation, bowel wall thickening, mesenteric edema, vascular engorgement on CT	anticoagulation
LE	0.2–6.0%	severe abdominal pain, diarrhea, nausea, vomiting, ascites	elevated CRP and lactic acid levels, leukopenia, elevated IgA levels, AutoABs	abdominal CT showing target sign (wall thickening above 3 mm), comb sign (engorgement of the mesenteric vessels), and increased attenuation of mesenteric fat; ascites	glucocorticoid pulse therapy and immunosuppressants with supportive measures, including antibiotics and fluid therapy; biologic agents such as rituximab or belimumab
IPO	approx. 1.7%	symptoms suggestive of obstruction – vomiting and constipation, preceded by diarrhea and weight loss	AutoABs (including anti-Ro and anti-RNP antibodies) and histopathological findings	air-fluid levels and dilated intestinal loops on CT/X-ray radiography; co-occurrence of bilateral hydronephrosis (CT/USG)	glucocorticoids and immunosuppressants with supportive measures including decompression or surgery
LP	isolated – single cases; serositis with pleural/pericardial effusion – 10–20%, although peritonitis remains rare in these cases	chronic – exudative ascites; acute – nonspecific symptoms, including nausea, vomiting, abdominal pain, and tenderness	peritoneal fluid analysis combined with AutoABs – to exclude infection or malignancy	ascites on CT	no standardized guidelines, good results of methylprednisolone pulse therapy with possible addition of immunosuppressants
Oral manifestations	approx. 30%	oral ulcers, hyperkeratosis, erosion, oral pigmentation, fissured tongue	haplotype C of CD34 gene polymorphism, VEGF pathway	–	proper dental hygiene, HCQ, glucocorticoids, immunosuppressants
Lupus pancreatitis	0.2–8.2% (0.4–1.1 cases per 1,000 patients)	acute abdominal pain located in the epigastric region, radiating to the back, nausea and vomiting	elevated amylase and lipase levels, high anti-dsDNA antibody levels, low complement (C3/C4) levels	AP on CT	high-dose corticosteroids (methylprednisolone pulse therapy), CPA mycophenolate mofetil, HCQ
LH	3–8%	fatigue, yellowing of skin/sclera, dark urine/pale stools, ascites, skin rashes, photosensitivity	elevated ALT/AST, LDH, total/direct bilirubin, total bile acid levels, decreased haptoglobin, platelet count, and RBC count	hepatomegaly, cholecystitis, ascites on CT	corticosteroids (methylprednisolone pulse therapy, oral prednisone/prednisolone), azathioprine

SLE – systemic lupus erythematosus; GERD – gastroesophageal reflux disease; PLE – protein-losing enteropathy; IPO – intestinal pseudo-obstruction; LE – lupus enteritis; LP – lupus peronitis; LH – lupus hepatitis; AP – acute pancreatitis; CRP – C-reactive protein; IgA – immunoglobulin A; ALT – alanine aminotransferase; AST – aspartate aminotransferase; LDH – lactate dehydrogenase; RBCs – red blood cells; AutoABs – autoantibodies; anti-dsDNA – anti-double-stranded DNA antibodies; C3/C4 – complement components 3 and 4; D-dimer – fibrin degradation product; VEGF – vascular endothelial growth factor; CD34 – cluster of differentiation 34; CT – computed tomography; USG – ultrasonography; ^{99m}Tc-HSA – technetium-99m labelled human serum albumin; HCQ – hydroxychloroquine; CPA – cyclophosphamide.

Materials and methods

The literature search was conducted using the PubMed, Embase, and Google Scholar databases, as well as references from relevant articles and online sources. The following medical keywords were used, the majority of which were Medical Subject Headings (MeSH) terms: "lupus" OR "SLE" combined with "GI system", "colitis", "mesenteric vasculitis", "intestinal pseudo-obstruction (IPO)", "GERD", "gastroesophageal reflux", "oral cavity", "peritonitis", "hepatitis", "pancreatitis", "enteritis", "mesenteric vascular thrombosis", "lupus-associated protein-losing enteropathy (LUPLE)", "glucocorticosteroids", "hydroxychloroquine", "azathioprine (AZA)", "nonsteroidal anti-inflammatory drugs (NSAIDs)", "cyclophosphamide", and "belimumab". The authors screened article titles and abstracts to identify relevant publications. Original articles and review papers published in 2020 or later were included. Older articles and case reports were considered only when more recent literature was unavailable. The final literature search was performed on May 25, 2025. The exclusion criteria were as follows: articles published before 2020, publications in languages other than English, studies with very small sample sizes, and case reports when review articles or meta-analyses were available.

Oral cavity

Oral cavity involvement is a common manifestation of SLE. Du et al. reported that the prevalence of oral manifestations may reach 30%.[7] Oral ulcers, the most frequently observed oral pathology, are included in the European Alliance of Associations for Rheumatology (EULAR) recommendations for the management of SLE. These lesions may be painful or remain asymptomatic and unnoticed by the patient. Clinically, they typically present as pale yellow or grayish-white superficial ulcers.[8]

Other reported oral manifestations include hyperkeratosis, erosions, oral pigmentation, fissured tongue, and xerostomia resulting from oral dryness.[9] The hard palate and buccal mucosa are the most common sites of involvement. Therefore, proper oral and dental hygiene represents an important aspect of prophylaxis and supportive care. Management of oral lesions includes a variety of therapeutic approaches, with antimalarial agents considered first-line treatment. In addition, standard SLE therapies, including glucocorticosteroids, immunosuppressive agents, and other systemic treatments, may also be used.[10]

GERD

One study reported that nearly half of patients with SLE experienced esophageal involvement, most commonly regurgitation (27.3%) and heartburn (24.6%).[11] The pathophysiology of esophageal motility disorders in SLE remains unclear. Earlier studies suggested that these abnormalities may result from ischemic vasculitis, muscle atrophy, or inflammatory involvement of the esophageal musculature. Symptomatic treatment is primarily based on proton pump inhibitors, prokinetic agents, and high-dose H_2-receptor blockers.[6]

Intestinal pseudo-obstruction

Intestinal pseudo-obstruction is a rare gastrointestinal manifestation of SLE, characterized by symptoms resembling mechanical or functional bowel obstruction despite the absence of an identifiable obstructive cause.[12] It may occur either as the initial manifestation of the systemic disease or as one of multiple clinical manifestations of SLE.[13] The prevalence of IPO is estimated to affect approx. 1–2% of patients with SLE.[14] Several theories have been proposed regarding the pathophysiology of this manifestation of lupus; however, none has been conclusively confirmed due to the limited number of reported cases. Based on current evidence, 2 principal mechanisms have been suggested. One hypothesis involves vasculitis affecting the smooth muscles of the GI tract, leading to ischemia and dysfunction of the affected tissue, which subsequently impairs normal smooth muscle contraction. The 2nd hypothesis, frequently associated with ureterohydronephrosis and other genitourinary complications, suggests generalized smooth muscle dysmotility caused by myopathic or neurogenic abnormalities. These alterations are thought to develop as a result of autoantibodies and immune complexes directed against smooth muscle tissue and its vasculature (Fig. 2).[12]

Patients typically present with symptoms suggestive of bowel obstruction, including nausea, vomiting, and constipation; however, episodes of diarrhea may precede the acute phase of the disease. In addition, substantial weight loss is often observed before the onset of acute symptoms.[15] In more severe cases, patients may also present with fever.[16] Physical examination may reveal absent bowel sounds on auscultation, as well as diffuse abdominal tenderness and muscular guarding.[17] Laboratory findings are similar to those observed in lupus enteritis (LE), including elevated C-reactive protein (CRP) levels and the presence of SLE-associated autoantibodies.[18] Diagnostic imaging typically demonstrates features of bowel obstruction, such as air–fluid levels on abdominal radiography and dilated intestinal loops. A characteristic finding in this manifestation is the coexistence of bilateral hydronephrosis, which is frequently identified on computed tomography (CT).[19] This phenomenon is considered highly prevalent, and its absence is often regarded as atypical and highlighted in case reports. Management of this manifestation of lupus includes not only corticosteroids and immunosuppressive agents, but also supportive measures such as nasogastric decompression, electrolyte correction, and, in selected cases, surgical intervention. Standard pharmacological treatment most commonly involves pulse methylprednisolone therapy, cyclophosphamide (CPA), and hydroxychloroquine (HCQ).[17]

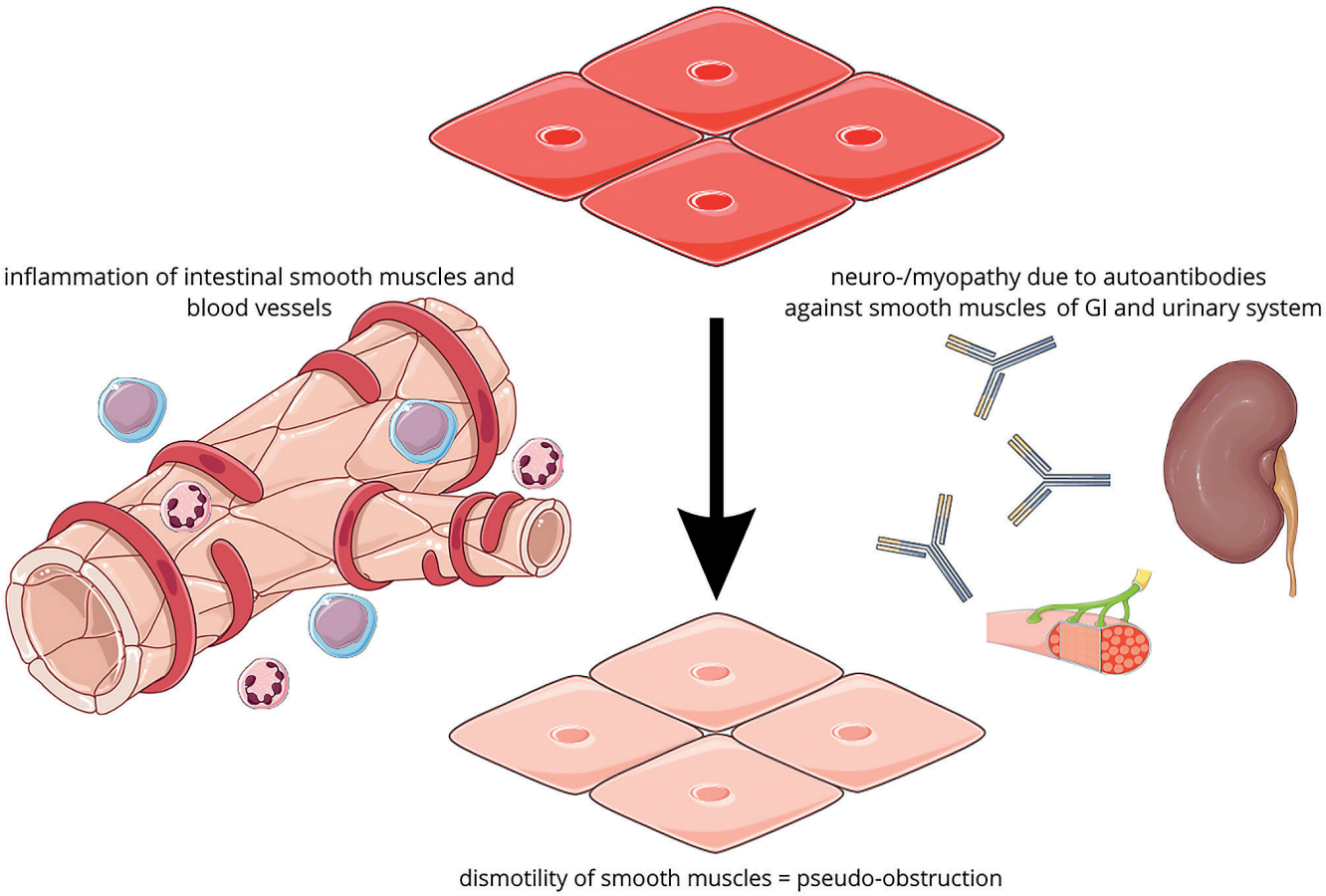

Fig. 2. Graphical representation of the pathophysiology of lupus pseudo-obstruction

Lupus peritonitis

Lupus peritonitis (LP) is one of the least frequent GI manifestations of SLE, and it is not described in the literature as often as other lupus features, such as lupus nephritis or enteritis. Serositis, defined as inflammation of serous membranes, is a well-established criterion of SLE. It typically involves the pleura or pericardium, with peritonitis being a far less common presentation. The reported prevalence of serositis in SLE ranges from 10% to 20%, depending on the cohort studied.[20] Isolated LP is particularly rare, but cases in which it was the first manifestation of SLE have been documented.[21] The pathophysiology of LP is believed to involve the formation of pathological immune complexes in the peritoneum and inflammation of local vessels and serosal surfaces. Clinically, LP can present in an acute or chronic form. The former is associated with nonspecific acute abdominal symptoms, such as abdominal pain, nausea, vomiting, and tenderness.[22] Chronic manifestation is characterized by exudative ascites, which should be differentiated from appendicitis and neoplastic or infectious causes.

Diagnostic evaluation should include autoantibody testing (e.g., ANA, anti-dsDNA), complement levels, and peritoneal fluid analysis to differentiate LP from other causes.[23] Currently, no specific guidelines for the treatment of LP have been developed due to its rarity and the limited amount of available literature. Despite this, available case reports demonstrate high effectiveness of intravenous or oral pulse methylprednisolone therapy.[22] Immunosuppressants, such as mycophenolate mofetil, are sometimes administered together with glucocorticosteroids with satisfactory results.

Lupus hepatitis

Lupus hepatitis (LH) is a liver dysfunction associated with SLE. Since the liver is rarely a primary target organ in SLE, liver function abnormalities are not included in the diagnostic or classification criteria for the disease.[24] However, up to 50% of patients with SLE are reported to experience abnormal liver function at some point in their lives, possibly related to SLE. This may result from LH, coexisting autoimmune liver diseases, or liver damage caused by non-autoimmune conditions.[25] Lupus hepatitis occurs in approx. 3–8% of affected individuals and is characterized by mildly elevated transaminase levels, often without clinical symptoms.[26] Elevations may be observed in liver enzymes, including alanine aminotransferase (ALT), aspartate aminotransferase (AST), alkaline phosphatase (ALP), and gamma-glutamyl transferase (GGT), as well as bilirubin, inflammatory markers such as erythrocyte

sedimentation rate (ESR) and CRP, and SLE-specific serological markers. Laboratory findings may also reveal reduced complement levels, including C3, C4, and C1q.[27]

Strongly positive complement C1q deposits observed in liver immunohistochemistry are highly indicative of LH, whereas such deposits are absent in patients with other hepatic disorders.[28] Patients with LH characteristically exhibit elevated anti-ribosomal P antibody levels, and liver biopsy usually reveals lobular inflammation. These features may be valuable in differentiating LH from other chronic hepatic pathologies observed in patients with SLE.[29] Several studies have reported that elevated liver enzyme levels in patients with SLE usually normalize following corticosteroid treatment. In contrast, a diminished initial therapeutic response to corticosteroid treatment has been observed in patients with autoimmune hepatitis (AIH).[30]

Autoimmune hepatitis is a chronic, immune-mediated inflammatory liver disease characterized by elevated levels of circulating autoantibodies, hypergammaglobulinemia, and increased serum transaminase activity.[31] The overlap syndrome involving AIH and SLE represents a rare complication.[32] Differential diagnosis between AIH and SLE-associated LH can be demanding. Certain parameters, such as elevated antinuclear antibody (ANA) and immunoglobulin (IgG) levels, are commonly observed in both conditions. However, markers more specific for AIH and typically absent in SLE include antibodies against soluble liver antigen (SLA), liver–pancreas antigen, smooth muscle antibodies (SMA) with specificity for F-actin, and microsomal autoantibodies such as anti-liver kidney microsomal antibodies (anti-LKM antibodies).[33,34] A key distinguishing factor between AIH and SLE is liver histopathology. Patients with AIH typically present with characteristic biopsy findings, including interface hepatitis, hepatocyte rosetting, emperipolesis, and fibrosis. In comparison, liver histology in SLE usually reveals fatty degeneration or hydropic changes in hepatocytes, which are typically associated with drug-induced toxicity or nonspecific hepatic involvement.[35]

A recent systematic review summarized findings from 4 studies investigating the treatment of lupus-associated hepatitis, including SLE–autoimmune hepatitis (SLE-AIH) overlap, SLE-AIH–primary biliary cholangitis (PBC) overlap, and LH, encompassing a total of 59 patients. Overall, approx. 2/3 of patients (39/59) experienced favorable outcomes, while 11 patients relapsed and 3 died. Intravenous methylprednisolone (IVMP) and oral prednisolone (0.5–1 mg/kg/day), often in combination with AZA, formed the basis of both induction and maintenance therapy. Although higher doses of prednisolone (>0.25 mg/kg/day) may correlate with improved response rates, the available evidence remains limited and of low certainty. Conservative management may be appropriate for patients with mild disease presentations.

Differentiating SLE-associated hepatitis from AIH remains clinically challenging due to overlapping features; however, specific serological and histological markers may assist in establishing the diagnosis. Current limited evidence suggests that immunosuppressive therapy with corticosteroids and AZA is generally effective, although high-quality controlled studies are needed to establish standardized treatment approaches.[10]

Acute pancreatitis

Acute pancreatitis (AP) is an acute inflammatory condition of the pancreas that constitutes a gastroenterological emergency.[36] Pancreatitis is primarily caused by alcohol consumption and gallstones. Pancreatic involvement in SLE remains a rare occurrence.[37] Pancreatitis associated with SLE has a reported prevalence ranging from 0.2% to 8.2% and an annual incidence of 0.4–1.1 cases per 1,000 patients.[38] If not treated promptly, more than half of patients with SLE-related AP are at risk of developing complications. The onset of pancreatitis in SLE is multifactorial and involves a complex pathophysiology, including immune-mediated pancreatic damage and vasculitis.[39] Yuan et al. identified 132 cases of abdominal pain considered to be SLE-related. Among these, lupus-associated pancreatitis accounted for 17.4% of patients (23 out of 132).[40] One study reported that in 97% of cases, the diagnosis of pancreatitis was confirmed by laboratory evidence of elevated serum amylase or lipase levels.[41] Pancreatic involvement represents a severe and potentially life-threatening clinical condition, often occurring as part of widespread multiorgan involvement, such as macrophage activation syndrome (MAS).[42] Recent studies suggest that favorable outcomes in the management of acute pancreatitis associated with SLE (SLEAP) may be achieved with higher doses of glucocorticoids and intensified immunosuppressive therapy. In cases of chronic pancreatitis occurring in the context of SLE, limited evidence supports the use of induction therapy with prednisolone (20–60 mg/day), often in combination with other immunomodulatory agents such as CPA. However, the lack of high-quality randomized controlled trials (RCTs) underscores the need for further clinical research to inform evidence-based therapeutic guidelines. Further clinical studies are essential to establish clear and effective treatment strategies for both acute and chronic SLE-associated pancreatitis.[10]

Pancreatitis may present as the first manifestation of previously undiagnosed SLE or occur during the course of the disease. It is commonly associated with high disease activity, which increases the risk of severe outcomes. Patients experiencing severe disease flares are particularly vulnerable to developing fatal complications due to pancreatitis; therefore, appropriate therapeutic interventions are required. However, standardized guidelines regarding the type, dosage, and duration of corticosteroid therapy for SLE-associated pancreatitis are currently lacking, highlighting the need for further research in this area.

Lupus enteritis

Lupus enteritis is one of the most common specific GI manifestations of SLE. It may be either the first manifestation of the systemic disease or may develop concurrently with other symptoms.[43] It may also arise because of suboptimal or inappropriate management of lupus.[44] The pathophysiology of LE is not fully understood. It is suggested that deposition of immune complexes, as well as abnormal complement activation, leads to inflammation and microvascular damage, causing edema and thrombosis of intestinal blood vessels. Consequently, ischemia may lead to ulceration or even perforation.[45] The symptoms accompanying LE are usually nonspecific. They include severe abdominal pain – most commonly diffuse, although some reports indicate that it may be localized – abdominal distention, diarrhea without blood or mucus, nausea, dry heaving, and vomiting.[46] They present with varying degrees of severity, ranging from moderate to severe, and may be accompanied by symptoms involving other affected systems, such as joint pain, dysuria or anuria, and uterine bleeding.[47] Physical examination may reveal sluggish or completely absent bowel sounds, as well as abdominal tenderness or muscle guarding. However, physical examination findings may also be completely normal.

Most often, LE affects the jejunum and ileum. It can also involve different parts of the colon and, very rarely, the rectum.[48] Laboratory findings can be divided into nonspecific and lupus-associated abnormalities. The former include elevated CRP levels (ranging from low elevations to approx. 100 mg/L), erythrocyte sedimentation rate (ESR), and lactate dehydrogenase (LDH).[49] The latter include positive ANA, anti-Ro, anti-Sm, anti-dsDNA, anti-La, and anti-U1 RNP antibodies. Decreased complement protein levels (C3 and C4) are also frequently observed.[50] Blood morphology findings are nonspecific; however, low hemoglobin (Hb) levels are common, often due to blood loss, whereas white blood cell (WBC) counts may be either elevated or decreased.[51]

The gold standard imaging modality for diagnosis is abdominal CT. Possible findings include edema, thickening of the intestinal wall, and dilated intestinal loops.[47] There are no pathognomonic signs of LE; however, 3 specific findings strongly raise suspicion of this disease when present together: the target sign (bowel wall thickening above 3 mm), the comb sign (engorgement of mesenteric vessels), and increased attenuation of mesenteric fat.[47] Mild ascites is also frequently observed on CT. Pathological changes in the thorax, such as pleural effusion, may also be present.[52] Other diagnostic methods include ultrasonography.[53] Although it is easier to perform, its diagnostic accuracy is limited. In contrast, endoscopic procedures with biopsy, despite being more invasive, frequently demonstrate nonspecific inflammatory findings.[54] First-line treatment of LE includes glucocorticosteroids, mainly high-dose intravenous methylprednisolone administered in pulses (200 mg, 300 mg, or 500 mg daily) for several days, followed by oral prednisone therapy. Most patients improve rapidly; however, treatment is often supplemented with 1 or 2 pulses of CPA and daily HCQ therapy. Mycophenolate mofetil is sometimes used and has demonstrated efficacy in certain cases.[47,55] Supportive treatment with intravenous fluids, antibiotics, or vitamin supplementation also plays a vital role due to impaired GI tract function.[56] In more severe cases, this treatment may be insufficient. Biologic agents such as rituximab or belimumab are used in these situations, generally with favorable therapeutic outcomes.[10,52]

Lupus-associated protein-losing enteropathy

Protein-losing enteropathy is a rare manifestation of SLE characterized by excessive protein loss through the GI tract.[57] It leads to hypoproteinemia, hypoalbuminemia, and lymphopenia. Hypoalbuminemia may result in ascites, pleural effusion, pericardial effusion, and generalized edema.[58,59] In the pathogenesis of PLE, a key role is played by increased vascular permeability due to vessel damage caused by cytokine activity and complement deposition, as well as by vasculitis and intestinal lymphangiectasia.[58]

The early stage of PLE is frequently asymptomatic and difficult to detect. The most common symptom, with a prevalence of up to 89.7%, is peripheral edema. Among GI manifestations, the most frequent are abdominal distension, diarrhea, abdominal pain, and vomiting.[60] Protein-losing enteropathy should be considered in patients with hypoalbuminemia in the absence of significant proteinuria or malnutrition and with positive ANA results. Relevant diagnostic tests include magnetic resonance (MR) lymphangiography, alpha-1-antitrypsin clearance, and technetium-99m human serum albumin scintigraphy (^{99m}Tc-HSA); however, in low- and middle-income countries, PLE is often diagnosed by exclusion.[61] Technetium-99m human serum albumin scintigraphy is important because it can localize the site of protein leakage.[62] Differential diagnoses include right-sided heart failure, inflammatory bowel disease (IBD), chronic liver diseases, renal conditions associated with significant proteinuria, malnutrition, malabsorption syndrome, amyloidosis, malignancies, nonsteroidal anti-inflammatory drug-induced enteropathy, and post-chemotherapy protein loss.[58] The condition is treated with glucocorticoids and immunosuppressants such as CPA, mycophenolate mofetil, and cyclosporine A.[60]

Lupus mesenteric vasculitis

Mesenteric vasculitis is an uncommon GI complication of SLE that should be considered in patients presenting with abdominal pain. Lupus mesenteric vasculitis (LMV) is reported to affect between 0.2% and 9.7% of individuals with SLE, with a higher prevalence in Asia. It can be

the initial manifestation of SLE and occurs more frequently in pediatric patients.[63] The superior mesenteric artery is affected in 80–85% of cases, with the ileum and jejunum being more commonly involved than the large intestine and rectum. Patients with high disease activity are at a greater risk of developing LMV compared to those with inactive SLE.[64] The pathogenesis of this condition is associated with the deposition of immune complexes in vessel walls, as well as thrombosis resulting from the presence of antiphospholipid antibodies.[65] Case reports suggest that LMV may develop following GI infections; however, a definitive association has not been fully confirmed.[63] The symptoms of LMV result from ischemia, which may progress to organ infarction. The most common clinical manifestations include abdominal pain, tenderness, and rectal bleeding, while nausea, vomiting, and diarrhea are also frequently reported.[66,67] Thrombotic occlusion typically leads to acute, sudden-onset abdominal pain, whereas chronic ischemia presents with postprandial pain and progressive weight loss. Urinary symptoms, such as lupus cystitis and dysuria, are associated with LMV in approx. 22.7% of cases.[68]

A prompt diagnosis is essential to prevent complications, but it remains highly challenging. The imaging modality of choice is contrast-enhanced CT of the abdomen, which is considered the gold standard.[68] Typical radiological findings include bowel wall thickening, bowel loop dilation, intestinal wall enhancement known as the "target sign", vascular occlusions, and ascites. Characteristic signs such as the "fence-like" pattern and the double halo sign may also be observed.[69] The differential diagnosis should include infections, malignancies, adverse drug reactions, and atherosclerotic mesenteric ischemia. Management of the underlying disease is a key aspect of treatment.[65] Anti-inflammatory and immunosuppressive therapy includes glucocorticoids, CPA, AZA, and mycophenolate mofetil.[10] In SLE-related LMV, high-dose intravenous glucocorticosteroids are the first-line treatment,[68] while CPA is reserved for refractory cases. Biologic agents, such as rituximab and tumor necrosis factor alpha (TNF-α) inhibitors, have also shown efficacy. Given the risk of perforation, surgical intervention, including segmental bowel resection, should be considered at an early stage.[70]

Mesenteric vascular thrombosis in SLE

Mesenteric vascular thrombosis is rare in SLE, and its frequency is difficult to determine due to the limited number of studies. The risk of thrombosis increases in the presence of concurrent vasculitis and antiphospholipid syndrome (APS).[71] Antiphospholipid antibodies are present in approx. 20–40% of patients with SLE, and within 20 years, 50–70% of these individuals are likely to develop APS.[72] Patients with SLE can develop both arterial and venous mesenteric thrombosis.[73] This condition can lead to acute intestinal ischemia, necrosis, perforation, and hemorrhage. Clinically, it may present with acute abdominal pain and vomiting. Laboratory tests often reveal elevated WBC count, D-dimer, CRP, and lactic acid levels.[74] Computed tomography findings may include bowel dilatation, bowel wall thickening, mesenteric edema, or vascular engorgement. Thrombi are often clearly visualized within mesenteric vessels on imaging.[75] However, definitive diagnosis often requires exploratory laparotomy. Bowel segments affected by transmural infarction must be resected, and in some cases, a second-look laparotomy may be necessary.[76] Current evidence suggests that anticoagulation should be the primary conservative treatment for mesenteric vein thrombosis, as it is associated with lower mortality, fewer complications, and improved outcomes compared to other therapies.[77] Mesenteric thrombosis is a rare but potentially life-threatening cause of abdominal pain, with a mortality rate of up to 50% if left untreated.[78]

Drugs side effects

Gastrointestinal manifestations are common in patients with SLE, but most are related to adverse drug reactions and infections (Table 2).[13] Glucocorticosteroids possess anti-inflammatory and immunosuppressive properties, which is why they are widely used in the treatment of autoimmune diseases such as SLE. Dexamethasone belongs to this group and has been associated with the occurrence of AP. Clinical cases have been reported in which AP developed after re-administration of dexamethasone.[79] There is an increased risk of AP in patients receiving at least 1 of the following glucocorticosteroids: cortisol, dexamethasone, or prednisolone.[80] Increasing the steroid dose may further elevate the risk of AP. Acute pancreatitis usually develops within 4–14 days after the initial exposure.

The mechanism by which AP occurs during glucocorticosteroid therapy remains unclear. However, it may be related to changes in lipid and calcium metabolism resulting from the systemic effects of corticosteroids. Another mechanism identified in animal studies suggests that corticosteroids may obstruct small pancreatic ducts, leading to increased viscosity of pancreatic secretions, which in turn causes pancreatic changes such as vacuolization of acinar cells, necrosis of peripancreatic fat, and hyperplasia of the islets of Langerhans.[81]

Hydroxychloroquine and, to a lesser extent, chloroquine (CQ) are commonly used in rheumatic autoimmune disorders (RADs), such as SLE. The most common side effects observed in RAD patients receiving HCQ or CQ are GI disturbances.[82] These include decreased appetite, nausea, vomiting, diarrhea, bloating, and a burning retrosternal sensation. The most common adverse effect is GI upset, which may be severe enough to cause loss of appetite. It occurs more frequently with CQ (20%) than with HCQ (10%).[83] These manifestations are usually transient and resolve over time or after dose reduction.[83]

Tacrolimus is a calcineurin inhibitor mainly metabolized by CYP3A. Ritonavir is a CYP3A and P-gp inhibitor, which

Table 2. Drugs used in the treatment of SLE associated with gastroenterological complications

Drug side effects	Prevalence of each manifestation	Common symptoms	Diagnostic markers	Imaging findings	Standard treatments
Glucocorticosteroids	–	AP	elevated amylase and lipase levels	vacuolation of acinar cells, necrosis of peripancreatic fat, and hyperplasia of the islets of Langerhans	drug withdrawal
HCQ, CQ	CQ (20%), HCQ (10%)	decreased appetite, nausea, vomiting, diarrhea, bloating and burning sensation in the retrosternal area	–	–	these manifestations are transient and resolve over time or after dose reduction
Tacrolimus	–	nausea, vomiting, fatigue, weakness, loss of appetite, abdominal pain, slowed speech, and peripheral neuropathy	blood tacrolimus levels increased above the toxic range	paralytic ileus was diagnosed on abdominal X-ray imaging, resulting from drug–drug interaction between tacrolimus and nirmatrelvir/ritonavir	tacrolimus was discontinued and nirmatrelvir/ritonavir was continued until completion of treatment
Leflunomide	10–20%, colitis – very rare, only a few cases reported in the literature	nausea, abdominal pain, diarrhea, leflunomide-induced colitis	diagnosis is challenging because symptoms may appear at different times, even several years after treatment initiation	abdominal CT revealed colitis	drug withdrawal
NSAIDs	common, approx. 25% of patients without classical risk factors still developed ulcers during high-dose NSAID therapy	gastric and duodenal ulcers	epigastric pain, endoscopic examination	erosive and ulcerative lesions commonly occurring in the gastric antrum, bleeding and subepithelial hemorrhages on endoscopic examination	drug withdrawal, proton pump inhibitors, H_2 blockers
Azathioprine	3.25%; however, clinical experience suggests that the incidence may be at least twice as high	AP	elevated amylase and lipase levels	CT showed evidence of AP tests exclude other common causes of pancreatitis	drug withdrawal
CPA	–	vomiting, diarrhea, nausea, stomatitis	–	–	resolution of short-term toxicity
Belimumab	–	AP	elevated amylase and lipase levels	CT showed evidence of AP	drug withdrawal

SLE – systemic lupus erythematosus; NSAIDs – nonsteroidal anti-inflammatory drugs; HCQ – hydroxychloroquine; CQ – chloroquine; CT – computed tomography; CPA – cyclophosphamide..

may increase the plasma concentration of medications that are substrates of the CYP3A and P-gp enzymatic systems. Ritonavir is used in the treatment of HIV infection. For this reason, significant interactions between tacrolimus and nirmatrelvir/ritonavir have been reported. Blood tacrolimus concentrations increased above the therapeutic threshold in all patients. These patients presented with various symptoms, including nausea, vomiting, fatigue, weakness, loss of appetite, abdominal pain, slowed speech, and peripheral neuropathy, after taking nirmatrelvir/ritonavir without discontinuing tacrolimus.[84]

Leflunomide is an oral disease-modifying antirheumatic drug with anti-inflammatory and immunomodulatory properties.[85] Many experimental models and clinical studies have shown that leflunomide has a beneficial effect in SLE.[86] Common side effects of leflunomide include GI symptoms such as nausea, abdominal pain, and diarrhea, occurring in 10–20% of treated patients. More serious adverse effects, such as colitis, have also been reported. Leflunomide-induced colitis is a very rare adverse effect, with only a few cases described in the literature. The mechanism of this condition is not well understood. Diagnosis is challenging because symptoms may appear at different times, even several years after initiation of treatment. The median time to symptom relief following discontinuation of leflunomide was approx. 3 weeks, but ranged from 3 days to 7 weeks.[87]

Nonsteroidal anti-inflammatory drugs (NSAIDs) rank among the most frequently used medications worldwide. Evidence from certain studies indicates that the use of NSAIDs and aspirin among the elderly population reaches a prevalence of 24.7%.[88] NSAIDs are commonly administered for the management of pain and discomfort associated with chronic rheumatologic conditions.[89]

According to data from 2 large cohort studies – ESTHER (n = 7,737) and the UK Biobank (n = 213,598) – low-dose aspirin administration independently contributes to an increased risk of developing gastric and duodenal ulcers during the early phase of treatment initiation.[90] The risk ratios (RRs) for the development of gastric and duodenal ulcers were 1.82 [95% CI: 1.58–2.11] and 1.66 [1.36–2.04], respectively, in the UK Biobank study, and 2.83 [1.40–5.71] and 3.89 [1.46–10.42], respectively, in the ESTHER study. According to data from studies conducted in Spain, the mortality rate associated with NSAID use is 5.6%, corresponding to 15.3 deaths per 100,000 users.[91] In the endoscopic picture, erosive and ulcerative lesions commonly occur in the gastric antrum; however, the pathological process may affect the entire gastroduodenal tract. These lesions may manifest in both acute and chronic forms, often presenting with bleeding and subepithelial hemorrhages on endoscopic examination. Ulcers typically heal without the formation of fibrotic scarring or structural deformities.[88] Recent studies comparing the safety profile of celecoxib at a daily dose of 200 mg with other NSAIDs in the context of GI adverse effects have demonstrated that celecoxib exhibits a superior GI safety profile relative to other NSAIDs.[92] Similarly, a previously conducted meta-analysis found that patients treated with meloxicam (a partially cyclooxygenase-2 (COX-2) selective agent) demonstrated a decreased occurrence of GI adverse effects, including dyspepsia, gastric and duodenal ulcers, and perforations. These findings suggest that COX-2 selective inhibitors are associated with a reduced risk of GI toxicity compared to nonselective NSAIDs. However, certain studies have demonstrated that GI symptoms during treatment with the nonselective NSAID naproxen were not indicative of mucosal injury. Although the risk of ulcer development correlates with the number of risk factors present, approx. 25% of patients without any classical risk factors still developed ulcers during high-dose NSAID therapy. These findings suggest that reliance solely on risk factor assessment is inadequate for determining the need for gastroprotection. A proactive preventive strategy is therefore recommended to ensure that all patients prescribed NSAIDs receive appropriate gastroprotective measures. Studies confirm that gastric and duodenal ulcers are common adverse effects associated with NSAID use, particularly among patients with rheumatologic conditions. Although NSAIDs are effective in the management of pain and inflammation, their use must be approached with caution because of GI risks, and appropriate preventive measures are required to reduce serious complications.

Azathioprine is a member of the immunosuppressive thiopurine group and is used to treat autoimmune diseases such as SLE. Retrospective data suggest that approx. 3.25% of patients treated with thiopurines experience AP, although clinical experience suggests that the incidence may be at least twice as high.[93] The exact pathogenesis of AZA-induced AP is not yet known, but its idiosyncratic nature suggests an allergic mechanism. The association between AZA and AP appears to be dose-independent and does not correlate with myelosuppression, suggesting an etiology independent of thiopurine methyltransferase activity. A delayed type II or IV allergic reaction or an immunological genetic predisposition has been proposed, with the former hypothesis supported by the fact that rechallenge with AZA causes recurrence of symptoms.[94] Azathioprine was one of 3 drugs in the study for which evidence of causing AP was demonstrated in RCTs.[95] If a patient is admitted with a high clinical suspicion of drug-induced pancreatitis (DIP), AZA should be discontinued immediately while awaiting tests to exclude other common causes of pancreatitis.

Cyclophosphamide is widely prescribed for the treatment of autoimmune diseases such as SLE. Treatment regimens for patients with SLE often cause short-term toxicity manifested by GI symptoms (e.g., vomiting and diarrhea).[96] Reported GI side effects include dose-related nausea, stomatitis, and a single case of hemorrhagic colitis. One report described a patient with colitis accompanied by fatal small bowel enteritis and pancolitis associated with 4 weeks of CPA therapy.[97]

Belimumab is a fully human recombinant IgG monoclonal antibody that targets and inhibits soluble B-lymphocyte stimulator (BLyS). Consequently, it induces apoptosis and downregulation of circulating B-cell clones. Studies have shown that patients with SLE have significantly higher BLyS levels than healthy controls. There have been some reports of AP associated with belimumab use. These cases were reported in women aged 40–49 years who had been receiving the drug for 1–2 years.[87]

Discussion

Systemic lupus erythematosus encompasses a wide array of GI manifestations that present significant diagnostic and therapeutic challenges. Our scoping review identified 11 distinct GI presentations, ranging from common, low-severity conditions such as GERD and oral ulcers to high-risk acute complications including LE, mesenteric vasculitis, and IPO. Although GERD and oral ulcers affect up to 1/3 of patients with SLE and are generally managed with proton pump inhibitors and antimalarials, delayed recognition of these symptoms as part of the lupus spectrum may prolong the time to definitive diagnosis and initiation of immunosuppressive therapy. Conversely, LE, mesenteric vasculitis, and IPO, while rarer, with prevalence ranging from 0.2% to 9.7%, carry risks of ischemia, perforation, and acute abdomen that require prompt contrast-enhanced CT imaging and aggressive treatment with high-dose intravenous corticosteroids, often supplemented with CPA pulses.

Between these extremes lie intermediate manifestations such as LH and pancreatitis, occurring in approx. 3–8%

and 0.2–8.2% of patients, respectively. These conditions require careful differentiation from autoimmune overlap syndromes. Serological markers such as anti-ribosomal P antibodies, anti-LKM antibodies, and complement C1q deposits, alongside targeted histological evaluation, may help clinicians choose between intensified immunosuppression and conservative management.

A unifying pathophysiological mechanism underlies several of these GI manifestations: immune complex-mediated vascular injury. Deposition of antinuclear antibody complexes activates complement pathways, damages the endothelium, and promotes thrombosis, thereby contributing to mesenteric vasculitis, LE, IPO, and PLE. Recognition of this shared process emphasizes the importance of early and aggressive immunosuppressive intervention.

Despite advances in imaging techniques and treatment strategies, significant diagnostic and therapeutic gaps persist. Reliance on advanced CT imaging and histological confirmation limits rapid diagnosis in resource-constrained settings, and the absence of consensus diagnostic criteria for rare manifestations such as PLE or IPO contributes to underreporting and inconsistent management. Moreover, therapeutic protocols for many GI manifestations are mainly derived from case series and small cohort studies, with no standardized guidelines regarding the dosing, duration, or selection of immunosuppressive agents.

Adding further complexity, the immunosuppressants and supportive medications used to manage SLE carry their own GI risks. NSAIDs predispose even low-risk patients to peptic ulcers, corticosteroids may precipitate AP, and AZA poses an idiosyncratic risk of DIP These iatrogenic complications require vigilant pharmacovigilance, patient education, and proactive gastroprotective measures as integral components of SLE management.

Looking forward, future research must prioritize prospective, multicenter cohort studies to establish the true incidence, risk factors, and outcomes of GI manifestations of SLE. The development of noninvasive biomarkers, such as serum complement fragments or anti-endothelial antibodies, could improve early detection and monitoring, while AI-assisted imaging algorithms may enhance the identification of subtle vasculitic changes on CT and ultrasound. Finally, RCTs comparing specific immunosuppressive regimens for GI involvement are urgently needed to support the development of evidence-based standardized treatment protocols.

Limitations of the study

The heterogeneous nature of SLE makes it difficult to establish universal conclusions, since individual patient presentations can vary widely. Many of the referenced studies include small patient cohorts, which limits the generalizability of the findings and highlights the need for larger multicenter investigations. Additionally, the absence of standardized diagnostic criteria for certain GI manifestations complicates comparisons across studies, while the predominance of cross-sectional data leaves long-term disease progression and treatment efficacy insufficiently explored. Moreover, some treatment recommendations are based on case reports or small-scale studies rather than robust RCTs.

Conclusions

Gastrointestinal manifestations in SLE span a spectrum from mild, common symptoms to life-threatening acute complications. Early recognition of the shared mechanisms of immune complex-mediated vascular injury may facilitate timely and aggressive therapy. Standardized diagnostic criteria, accessible imaging, and consensus treatment protocols remain critical gaps. Addressing these issues through multicenter studies, biomarker development, and RCTs may improve outcomes and optimize care for patients with SLE and GI involvement.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Wojciech Bajurny https://orcid.org/0009-0009-3444-9627
Julia Grabowska https://orcid.org/0009-0009-0957-7520
Amelia Pielech https://orcid.org/0009-0009-8832-5187
Natalia Struzik https://orcid.org/0009-0001-6479-3869
Jakub A. Mastalerz https://orcid.org/0009-0007-1974-149X
Magdalena Szmyrka https://orcid.org/0000-0001-6885-6153

References

1. Hoi A, Igel T, Mok CC, Arnaud L. Systemic lupus erythematosus. *Lancet*. 2024;403(10441):2326–2338. doi:10.1016/S0140-6736(24)00398-2
2. Siegel CH, Sammaritano LR. Systemic lupus erythematosus: A review. *JAMA*. 2024;331(17):1480. doi:10.1001/jama.2024.2315
3. Tian J, Zhang D, Yao X, Huang Y, Lu Q. Global epidemiology of systemic lupus erythematosus: A comprehensive systematic analysis and modelling study. *Ann Rheum Dis*. 2023;82(3):351–356. doi:10.1136/ard-2022-223035
4. Barber MRW, Drenkard C, Falasinnu T, et al. Global epidemiology of systemic lupus erythematosus. *Nat Rev Rheumatol*. 2021;17(9):515–532. doi:10.1038/s41584-021-00668-1
5. Murali N, El Hayek SM. Abdominal pain mimics. *Emerg Med Clin North Am*. 2021;39(4):839–850. doi:10.1016/j.emc.2021.07.003
6. Alharbi S. Gastrointestinal manifestations in patients with systemic lupus erythematosus. *Open Access Rheumatol Res Rev*. 2022;14:243–253. doi:10.2147/OARRR.S384256
7. Du F, Qian W, Zhang X, Zhang L, Shang J. Prevalence of oral mucosal lesions in patients with systemic lupus erythematosus: A systematic review and meta-analysis. *BMC Oral Health*. 2023;23(1):1030. doi:10.1186/s12903-023-03783-5
8. Benli M, Batool F, Stutz C, Petit C, Jung S, Huck O. Orofacial manifestations and dental management of systemic lupus erythematosus: A review. *Oral Dis*. 2021;27(2):151–167. doi:10.1111/odi.13271
9. García-Ríos P, Pecci-Lloret MP, Oñate-Sánchez RE. Oral manifestations of systemic lupus erythematosus: A systematic review. *Int J Environ Res Public Health*. 2022;19(19):11910. doi:10.3390/ijerph191911910
10. Williamson L, Hao Y, Basnayake C, Oon S, Nikpour M. Systematic review of treatments for the gastrointestinal manifestations of systemic lupus erythematosus. *Semin Arthritis Rheum*. 2024;69:152567. doi:10.1016/j.semarthrit.2024.152567

11. Asfuroglu Kalkan E, Kalkan C, Barutcu S, et al. Prevalence and determinants of gastrointestinal manifestations in patients with selected rheumatologic diseases. *Turk J Gastroenterol*. 2022;33(7):576–586. doi:10.5152/tjg.2022.21780
12. Zhang L, Xu D, Yang H, et al. Clinical features, morbidity, and risk factors of intestinal pseudo-obstruction in systemic lupus erythematosus: A retrospective case-control study. *J Rheumatol*. 2016;43(3):559–564. doi:10.3899/jrheum.150074
13. Tian XP, Zhang X. Gastrointestinal involvement in systemic lupus erythematosus: Insight into pathogenesis, diagnosis and treatment. *World J Gastroenterol*. 2010;16(24):2971–2977. doi:10.3748/wjg.v16.i24.2971
14. Zhou L, Cai SZ, Dong LL. Recent advances in pathogenesis, diagnosis, and therapeutic approaches for digestive system involvement in systemic lupus erythematosus. *J Digest Dis*. 2024;25(7):410–423. doi:10.1111/1751-2980.13307
15. Escalante-Pérez S, Guerra-Zarama S, Chavarriaga-Restrepo A, Echeverri-García A, Márquez-Hernández J, Pinto-Peñaranda LF. Pseudo-obstruction, an infrequent lupus manifestation: A case report [in Spanish] *Rev Colomb Reumatol*. 2023;30(1):59–62. doi:10.1016/j.rcreu.2021.04.009
16. Lan F, Li T, Zhou L, et al. Intestinal pseudo-obstruction in systemic lupus erythematosus complicated by Castleman disease: A case report. *Ann Transl Med*. 2022;10(20):1148. doi:10.21037/atm-22-4461
17. Wen J, Chen W, Gao L, Qiu X, Lin G. Systemic lupus erythematosus simultaneously presenting with visceral muscle dysmotility syndrome and mechanical intestinal obstruction clinically relieved by surgery: A case report and literature review. *BMC Gastroenterol*. 2022;22(1):32. doi:10.1186/s12876-022-02105-3
18. Nyabera A, Elfishawi M, Cuevas F, Riaz F, Abrudescu A. Intestinal pseudo-obstruction as the initial clinical presentation in systemic lupus erythematosus: A rare and severe disorder. *Case Rep Gastrointest Med*. 2020;2020:8873917. doi:10.1155/2020/8873917
19. Ng CS, Kan SL, Lim AL. AB0856 Chronic intestinal pseudo-obstruction with hydronephrosis: A case report on such disabling and rare complication of lupus. *Ann Rheum Dis*. 2021;80:1451–1452. doi:10.1136/annrheumdis-2021-eular.2854
20. Kao WC, Hu YC, Lee JH, et al. Serositis as an indicator of poor prognosis in pediatric systemic lupus erythematosus. *Pediatr Rheumatol*. 2025;23(1):36. doi:10.1186/s12969-025-01084-5
21. Kheyri Z, Laripour A, Ala M. Peritonitis as the first presentation of systemic lupus erythematous: A case report. *J Med Case Rep*. 2021;15(1):611. doi:10.1186/s13256-021-03216-3
22. Kiriakopoulos E, Perez V, Hoelle R. The great imitator strikes again: A case of a lupus flare-up presenting like an acute abdomen. *HCA Healthc J Med*. 2020;1(1):35–37. doi:10.36518/2689-0216.1012
23. Kosugi S, Yoshida T, Yoshimoto N, Itoh H, Oya M. A case of new-onset systemic lupus erythematosus with serositis in a maintenance hemodialysis patient. *Clin Med Insights Case Rep*. 2021;14:11795476211056172. doi:10.1177/11795476211056172
24. Praharaj DL, Mallick B, Nath P, Panigrahi SC, Padhan P, Sahu N. Unusual presentation of systemic lupus erythematosus in a young male: A case report. *J Clin Exp Hepatol*. 2021;11(2):264–269. doi:10.1016/j.jceh.2020.05.009
25. Liu J, Shen T, Li L, et al. A systemic lupus erythematosus patient with persistent elevated conjugated bilirubin as the initial symptom: A case report. *Medicine (Baltimore)*. 2024;103(6):e36999. doi:10.1097/MD.0000000000036999
26. Piga M, Vacca A, Porru G, Cauli A, Mathieu A. Liver involvement in systemic lupus erythematosus: Incidence, clinical course and outcome of lupus hepatitis. *Clin Exp Rheumatol*. 2010;28(4):504–510. PMID:20609296.
27. Afzal W, Haghi M, Hasni S, Newman K. Lupus hepatitis, more than just elevated liver enzymes. *Scand J Rheumatol*. 2020;49(6):427–433. doi:10.1080/03009742.2020.1744712
28. Zheng RH, Wang JH, Wang SB, Chen J, Guan WM, Chen MH. Clinical and immunopathological features of patients with lupus hepatitis. *Chin Med J (Engl)*. 2013;126(2):260–266. PMID:23324274.
29. Koshy JM, John M. Autoimmune hepatitis: SLE overlap syndrome. *J Assoc Physicians India*. 2012;60:59–60. PMID:23547416.
30. Czaja AJ, Morshed SA, Parveen S, Nishioka M. Antibodies to single-stranded and double-stranded DNA in antinuclear antibody-positive type 1-autoimmune hepatitis. *Hepatology*. 1997;26(3):567–572. doi:10.1002/hep.510260306
31. Fallatah HI, Akbar HO. Autoimmune hepatitis as a unique form of an autoimmune liver disease: Immunological aspects and clinical overview. *Autoimmun Dis*. 2012;2012:312817. doi:10.1155/2012/312817
32. Ferdous N, Ghosh J, Islam MN. A case of systemic lupus erythematosus with severe jaundice. *J Med*. 2022;23(2):162–164. doi:10.3329/jom.v23i2.60634
33. Beisel C, Weiler-Normann C, Teufel A, Lohse AW. Association of autoimmune hepatitis and systemic lupus erythematodes: A case series and review of the literature. *World J Gastroenterol*. 2014;20(35):12662–12667. doi:10.3748/wjg.v20.i35.12662
34. Lohse AW, Mieli-Vergani G. Autoimmune hepatitis. *J Hepatol*. 2011;55(1):171–182. doi:10.1016/j.jhep.2010.12.012
35. Chowdhary VR, Crowson CS, Poterucha JJ, Moder KG. Liver involvement in systemic lupus erythematosus: Case review of 40 patients. *J Rheumatol*. 2008;35(11):2159–2164. doi:10.3899/jrheum.080336
36. Li Z, Xu D, Wang Z, et al. Gastrointestinal system involvement in systemic lupus erythematosus. *Lupus*. 2017;26(11):1127–1138. doi:10.1177/0961203317707825
37. Dima A, Balaban DV, Jurcut C, Jinga M. Systemic lupus erythematosus-related acute pancreatitis. *Lupus*. 2021;30(1):5–14. doi:10.1177/0961203320978515
38. Breuer GS, Baer A, Dahan D, Nesher G. Lupus-associated pancreatitis. *Autoimmun Rev*. 2006;5(5):314–318. doi:10.1016/j.autrev.2005.11.004
39. Yang Y, Ye Y, Liang L, et al. Systemic-lupus-erythematosus-related acute pancreatitis: A cohort from south China. *Clin Dev Immunol*. 2012;2012:568564. doi:10.1155/2012/568564
40. Yuan S, Lian F, Chen D, et al. Clinical features and associated factors of abdominal pain in systemic lupus erythematosus. *J Rheumatol*. 2013;40(12):2015–2022. doi:10.3899/jrheum.130492
41. Nesher G, Breuer GS, Temprano K, et al. Lupus-associated pancreatitis. *Semin Arthritis Rheum*. 2006;35(4):260–267. doi:10.1016/j.semarthrit.2005.08.003
42. Muhammed H, Jain A, Irfan M, et al. Clinical features, severity and outcome of acute pancreatitis in systemic lupus erythematosus. *Rheumatol Int*. 2022;42(8):1363–1371. doi:10.1007/s00296-021-04834-2
43. Preet K, Mittal BR, Singh H, Kumar R, Sekar A, Kurdia K. 18F-FDG PET/CT in lupus enteritis presenting with fever and abdominal pain. *Clin Nucl Med*. 2024;49(5):e211–e212. doi:10.1097/RLU.0000000000005153
44. Gan H, Wang F, Gan Y, Wen L. Rare case of lupus enteritis presenting as colorectum involvement: A case report and review of literature. *World J Clin Cases*. 2023;11(34):8176–8183. doi:10.12998/wjcc.v11.i34.8176
45. Matsuyama S, Fukuda A, Ohana M. Gastrointestinal: Lupus enteritis with duodenojejunal fistula causing intestinal obstruction and gastrointestinal perforation. *J Gastroenterol Hepatol*. 2024;39(2):220–221. doi:10.1111/jgh.16373
46. Caruso VA, Khan M, Groudan K, Duong N. S4716 lupus enteritis: An uncommon presentation of lupus diagnosis. *Am J Gastroenterol*. 2024;119(10 Suppl):S2983–S2984. doi:10.14309/01.ajg.0001048232.13398.77
47. Chaparro CA, Bernal-Macías S, Muñoz OM. Lupus enteritis as systemic lupus erythematosus main manifestation: Two case reports. *SAGE Open Med Case Rep*. 2024;12:2050313X241247433. doi:10.1177/2050313X241247433
48. Smith LW, Petri M. Lupus enteritis: An uncommon manifestation of systemic lupus erythematosus. *J Clin Rheumatol*. 2013;19(2):84–86. doi:10.1097/RHU.0b013e318284794e
49. Costa CJ, Nguyen MTT, Vaziri H. S2872 tract on fire: Lupus enteritis. *Am J Gastroenterol*. 2024;119(10 Suppl):S1977–S1977. doi:10.14309/01.ajg.0001040856.74871.a2
50. Muñoz-Urbano M, Sangle S, D'Cruz DP. Lupus enteritis: A narrative review. *Rheumatology (Oxford)*. 2024;63(6):1494–1501. doi:10.1093/rheumatology/kead689
51. Shaik MR, Shaikh NA, Huang Y, Pandey A, Wheeler E, Mikdashi J. Autoimmune lupus enteritis with pan-gastrointestinal involvement in an adult patient with systemic lupus erythematosus: Complete response to belimumab. *J Community Hosp Int Med Perspect*. 2023;13(4):17–20. doi:10.55729/2000-9666.1185

52. Hussein AH, Altamimi BL. Rapid and remarkable response to rituximab in a case of lupus enteritis: Challenging the norm. *Mediterr J Rheumatol*. 2024;35(1):187. doi:10.31138/mjr.080624.rar
53. Chandwar K, Sahoo RR, Wakhlu A, Wakhlu A. Pneumatosis intestinalis: A rare manifestation of lupus enteritis. *BMJ Case Rep*. 2022;15(1): e247779. doi:10.1136/bcr-2021-247779
54. Almutairi R, Alkhudair D, Aldei A. Lupus enteritis as the early manifestation of systemic lupus erythematosus successfully managed with belimumab: A case report. *Cureus*. 2025;17(1). doi:10.7759/cureus.76926
55. Thomas AR, Jayan A, Chopra M, et al. S2652 Severe lupus enteritis complicated by intractable gastrointestinal hemorrhage. *Am J Gastroenterol*. 2022;117(10 Suppl):e1753. doi:10.14309/01.ajg.0000867248.77716.2b
56. Naeem F, Noor MU, Batool S, Anwer Khan SE, Akmal M. An atypical initial manifestation of systemic lupus erythematosus: Lupus enteritis accompanied by intestinal pseudo-obstruction and bilateral hydronephroureter. *Cureus*. 2023;15(12):e50628. doi:10.7759/cureus.50628
57. Abu Jheasha A, Alsharif T, Alwahsh R, Abumunshar A, Al-Ardah R, Abuturki A. Protein-losing enteropathy as the first presentation of systemic lupus erythematosus the first case reported in Palestine with systemic review. *Ann Med Surg (Lond)*. 2024;86(12):7458–7464. doi:10.1097/MS9.0000000000002733
58. Bektaş M, Taş O, Ordu M. A case of systemic lupus erythematosus presenting with intestinal lymphangiectasia-associated protein-losing enteropathy accompanying hyperinflammation. *Int J Rheum Dis*. 2023;26(3):591–598. doi:10.1111/1756-185X.14541
59. Kojima M, Hanaoka H, Aoki K, Matsushita H, Ito H, Yamada H. Successful treatment of lupus protein-losing enteropathy with belimumab: A case report. *Modern Rheumatol Case Rep*. 2024;8(2):264–266. doi:10.1093/mrcr/rxae010
60. Peng L, Li Z, Xu D, et al. Characteristics and long-term outcomes of patients with lupus-related protein-losing enteropathy: A retrospective study. *Rheumatol Immunol Res*. 2020;1(1):47–52. doi:10.2478/rir-2020-0006
61. Abdalla E, Mohymeed N, Nail AMA, et al. Protein-losing enteropathy as the first presentation of systemic lupus erythematosus: A case report from Sudan. *Clin Case Rep*. 2023;11(5):e7314. doi:10.1002/ccr3.7314
62. Ng DM, Sek K, Nossent J. Protein-losing enteropathy as a rare manifestation of systemic lupus erythematosus. *BMJ Case Rep*. 2023;16(8): e256680. doi:10.1136/bcr-2023-256680
63. Alsolaimani R. Mesenteric vasculitis and urinary system involvement presenting as the initial manifestations of systemic lupus erythematosus treated successfully with glucocorticoids and rituximab: A case report. *Cureus*. 2022;14(11):e31474. doi:10.7759/cureus.31474
64. Alshehri A, Alhumaidi H, Asseri Y, Ahmed MK, Omer H. Mesenteric vasculitis as a rare initial presentation of systemic lupus erythematosus: A case report. *Saudi J Med Med Sci*. 2020;8(3):223. doi:10.4103/sjmms.sjmms_206_19
65. Gnanapandithan K, Sharma A. Mesenteric vasculitis. In: StatPearls. Treasure Island, USA: StatPearls Publishing; 2025:Bookshelf ID: NBK546610. http://www.ncbi.nlm.nih.gov/books/NBK546610.
66. Xie JJ, Ling GC, Jiang YB, Zhang JY. Lupus with initial mesenteric vasculitis. *Rheumatol Immunol Res*. 2023;4(2):102–103. doi:10.2478/rir-2023-0015
67. Wang H, Gao Q, Liao G, Ren S, You W. Clinico-laboratory features and associated factors of lupus mesenteric vasculitis. *Rheumatol Ther*. 2021;8(2):1031–1042. doi:10.1007/s40744-021-00323-x
68. Leone P, Prete M, Malerba E, et al. Lupus vasculitis: An overview. *Biomedicines*. 2021;9(11):1626. doi:10.3390/biomedicines9111626
69. Zhu J, Lai J, Liu X, et al. Clinical characteristics and prognosis of childhood-onset lupus mesenteric vasculitis as the initial presentation: A case-control study. *Arthritis Res Ther*. 2023;25(1):248. doi:10.1186/s13075-023-03237-x
70. Chen YL, Meng J, Li C. Intestinal perforation with systemic lupus erythematosus: A systematic review. *Medicine (Baltimore)*. 2023;102(31): e34415. doi:10.1097/MD.0000000000034415
71. Kwok TSH, Basso Dias A, Cino M, Khalili K. Antiphospholipid syndrome presenting with inferior mesenteric vein thrombosis and associated rectal edema in a patient with systemic lupus erythematosus. *Clin Rheumatol*. 2022;41(9):2911–2912. doi:10.1007/s10067-022-06207-x
72. Nagy N, Papp G, Gáspár-Kiss E, Diószegi Á, Tarr T. Changes in clinical manifestations and course of systemic lupus erythematosus and secondary antiphospholipid syndrome over three decades. *Biomedicines*. 2023;11(4):1218. doi:10.3390/biomedicines11041218
73. Larsen ML, Nørgaard L, Linge P, et al. Molecular mechanisms underlying thrombosis in systemic lupus erythematosus: A systematic review. *Semin Arthritis Rheum*. 2025;72:152707. doi:10.1016/j.semarthrit.2025.152707
74. Reintam Blaser A, Starkopf J, Björck M, et al. Diagnostic accuracy of biomarkers to detect acute mesenteric ischaemia in adult patients: A systematic review and meta-analysis. *World J Emerg Surg*. 2023;18(1):44. doi:10.1186/s13017-023-00512-9
75. Reintam Blaser A, Koitmäe M, Laisaar KT, et al. Radiological diagnosis of acute mesenteric ischemia in adult patients: A systematic review and meta-analysis. *Sci Rep*. 2025;15(1):9875. doi:10.1038/s41598-025-94846-w
76. Acosta S, Salim S. Management of acute mesenteric venous thrombosis: A systematic review of contemporary studies. *Scand J Surg*. 2021;110(2):123–129. doi:10.1177/1457496920969084
77. Wang L, Wang E, Liu F, et al. A systematic review and meta-analysis on endovascular treatment as an attractive alternative for acute superior mesenteric venous thrombosis. *Vascular*. 2022;30(2):331–340. doi:10.1177/1708538121991270
78. Bala M, Catena F, Kashuk J, et al. Acute mesenteric ischemia: Updated guidelines of the World Society of Emergency Surgery. *World J Emerg Surg*. 2022;17(1):54. doi:10.1186/s13017-022-00443-x
79. Paramythiotis D, Karlafti E, Veroplidou K, et al. Drug-induced acute pancreatitis in hospitalized COVID-19 patients. *Diagnostics (Basel)*. 2023;13(8):1398. doi:10.3390/diagnostics13081398
80. Nango D, Hirose Y, Goto M, Echizen H. Analysis of the association of administration of various glucocorticoids with development of acute pancreatitis using U.S. Food and Drug Administration adverse event reporting system (FAERS). *J Pharm Health Care Sci*. 2019;5(1):5. doi:10.1186/s40780-019-0134-6
81. Minupuri A, Patel R, Alam F, Rather M, Baba R. Steroid-induced pancreatitis: Establishing an accurate association poses a challenge. *Cureus*. 2020;12(8):e9589. doi:10.7759/cureus.9589
82. Nirk EL, Reggiori F, Mauthe M. Hydroxychloroquine in rheumatic autoimmune disorders and beyond. *EMBO Mol Med*. 2020;12(8):e12476. doi:10.15252/emmm.202012476
83. Hładyj E, Sikora M, Felis-Giemza A, Olesińska M. Antimalarials: Are they effective and safe in rheumatic diseases? *Reumatologia*. 2018;56(3):164–173. doi:10.5114/reum.2018.76904
84. Zhang W, Zhang X, Han J, et al. Case report: Paralytic ileus resulted from nirmatrelvir/ritonavir-tacrolimus drug-drug interaction in a systemic lupus erythematosus patient with COVID-19. *Front Pharmacol*. 2024;15:1389187. doi:10.3389/fphar.2024.1389187
85. Chis R, Gladman D, Vajpeyi R, Cino M. A198 Lefluonomide-induced colitis in immunosupressed patient with systemic lupus erythematosus and rheumatoid arthritis. *J Can Assoc Gastroenterol*. 2021;4(Suppl 1): 222–223. doi:10.1093/jcag/gwab002.196
86. Cao H, Rao Y, Liu L, et al. The efficacy and safety of leflunomide for the treatment of lupus nephritis in Chinese patients: Systematic review and meta-analysis. *PLoS One*. 2015;10(12):e0144548. doi:10.1371/journal.pone.0144548
87. Bazigh I, Asfour M, Muddassir S, Mughni S. Acute pancreatitis after the use of belimumab in a patient with systemic lupus erythematosus: Case report and review of literature. *Cureus*. 2022;14(2):e22540. doi:10.7759/cureus.22540
88. Bordin DS, Livzan MA, Gaus OV, Mozgovoi SI, Lanas A. Drug-associated gastropathy: Diagnostic criteria. *Diagnostics (Basel)*. 2023;13(13):2220. doi:10.3390/diagnostics13132220
89. Scarpignato C, Hunt RH. Nonsteroidal antiinflammatory drug-related injury to the gastrointestinal tract: Clinical picture, pathogenesis, and prevention. *Gastroenterol Clin North Am*. 2010;39(3):433–464. doi:10.1016/j.gtc.2010.08.010
90. Nguyen TNM, Sha S, Chen L, Holleczek B, Brenner H, Schöttker B. Strongly increased risk of gastric and duodenal ulcers among new users of low-dose aspirin: Results from two large cohorts with new-user design. *Aliment Pharmacol Ther*. 2022;56(2):251–262. doi:10.1111/apt.17050

91. Joo MK, Park CH, Kim JS, et al. Clinical Guidelines for Drug-Related Peptic Ulcer, 2020 Revised Edition. *Gut Liver*. 2020;14(6):707–726. doi:10.5009/gnl20246
92. Honvo G, Lengelé L, Alokail M, Al-Daghri N, Reginster JY, Bruyère O. Safety of anti-osteoarthritis medications: A systematic literature review of post-marketing surveillance studies. *Drugs*. 2025;85(4):505–555. doi:10.1007/s40265-025-02162-4
93. Teich N, Mohl W, Bokemeyer B, et al. Azathioprine-induced acute pancreatitis in patients with inflammatory bowel diseases: A prospective study on incidence and severity. *J Crohns Colitis*. 2016;10(1):61–68. doi:10.1093/ecco-jcc/jjv188
94. Gordon M, Grafton-Clarke C, Akobeng A, et al. Pancreatitis associated with azathioprine and 6-mercaptopurine use in Crohn's disease: A systematic review. *Frontline Gastroenterol*. 2021;12(5):423–436. doi:10.1136/flgastro-2020-101405
95. Saini J, Marino D, Badalov N, Vugelman M, Tenner S. Drug-induced acute pancreatitis: An evidence-based classification (revised). *Clin Transl Gastroenterol*. 2023;14(8):e00621. doi:10.14309/ctg.0000000000000621
96. Yamamoto N, Tsuchiya Y, Fukuda M, Niiro H, Hirota T. A case report of drug interactions between nirmatrelvir/ritonavir and tacrolimus in a patient with systemic lupus erythematosus. *Cureus*. 2024;16(1):e52506. doi:10.7759/cureus.52506
97. Yang LS, Cameron K, Papaluca T, et al. Cyclophosphamide-associated enteritis: A rare association with severe enteritis. *World J Gastroenterol*. 2016;22(39):8844. doi:10.3748/wjg.v22.i39.8844

Genetic biomarker screening for drug-resistant epilepsy in low- and middle-income countries (LMICs)

Dina Kalinina[1,2,A,C,D], Alimzhan Muxunov[1,2,A,E], Zhassulan Utebekov[2,E], Gaziz Kyrgyzbay[2,E], Darkhan Kimadiev[2,E], Guldana Zhumabaeva[2,E], Joseph Almazan[1,E,F], Antonio Sarria-Santamera[1,A,E,F]

[1] Department of Medicine, Nazarbayev University School of Medicine, Astana, Kazakhstan
[2] Epileptology Centre, Republican State Enterprise (RSE) Medical Centre Hospital of the President's Affairs Administration of the Republic of Kazakhstan, Astana, Kazakhstan

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(5):919–928

Address for correspondence
Dina Kalinina
E-mail: dina.kalinina@nu.edu.kz

Funding sources
This study was conducted under the scientific and technical program BR28512409, "Development of a comprehensive innovative technology for diagnostic selection and surgical treatment of focal epilepsy resistant to drug therapy," funded by the Ministry of Science and Higher Education of the Republic of Kazakhstan.

Conflict of interest
None declared

Received on July 8, 2025
Reviewed on August 11, 2025
Accepted on August 18, 2025

Published online on May 20, 2026

Abstract

Drug-resistant epilepsy (DRE) presents a major clinical and economic challenge, particularly in low- and middle-income countries (LMICs), where healthcare resources are limited and treatment gaps remain significant. Although epilepsy surgery remains the most effective intervention for eligible DRE patients, outcomes are variable, with success rates ranging from 30% to 70%. Emerging evidence suggests that genetic biomarkers can inform patient selection, predict surgical outcomes, and guide treatment planning. This review explores the potential of integrating genetic testing into presurgical evaluation protocols in LMICs. It examines the role of specific gene mutations in pharmacoresistance, seizure localization, and structural brain abnormalities, with a focus on improving surgical success rates and reducing unnecessary interventions. Incorporating genetic stratification into clinical decision-making could enhance cost-effectiveness, minimize the burden on healthcare systems, and support the development of personalized treatment pathways. Advancing genetic research and building capacity in precision neurology are essential steps toward improving DRE management in resource-constrained settings.

Key words: drug-resistant epilepsy, surgical outcomes, low- and middle-income countries, genetic testing, patient selection

Cite as
Kalinina D, Muxunov A, Utebekov Z, et al. Genetic biomarker screening for drug-resistant epilepsy in low- and middle-income countries (LMICs). *Adv Clin Exp Med*. 2026;35(5):919–928. doi:10.17219/acem/209664

DOI
10.17219/acem/209664

Highlights

- Genetic testing improves presurgical evaluation in drug-resistant epilepsy (DRE), particularly in MRI-negative cases.
- Genetic screening may support epilepsy surgery decision-making by identifying patients most likely to benefit from resection.
- Early genetic diagnosis in DRE can reduce healthcare costs and influence treatment strategies.
- Implementation of genetic testing in low- and middle-income countries (LMICs) is limited by infrastructure, funding, and genetic literacy barriers.
- Targeted genetic testing strategies and international collaboration may improve DRE management and epilepsy care in LMICs.

Introduction

Epilepsy is one of the most common neurological disorders, affecting approx. 50 million individuals worldwide.[1,2] The condition manifests as recurrent unprovoked seizures, which are diagnosed after 2 unprovoked episodes or following a single seizure with a high risk of recurrence.[3] The World Health Organization's Global Action Plan 2022–2031[4] emphasizes the urgent need for improved epilepsy management, particularly in low- and middle-income countries (LMICs), where approx. 80% of people with epilepsy reside.[5]

Drug-resistant epilepsy (DRE) is defined by the failure of adequate trials of at least 2 appropriately selected and tolerated antiepileptic drugs to achieve sustained seizure control and presents significant challenges.[6] Drug-resistant epilepsy is correlated with increased morbidity, diminished quality of life, and elevated healthcare costs.[7] Although epilepsy surgery is the most effective treatment option for selected DRE cases, surgical outcomes vary considerably, with seizure freedom rates ranging from 30% to 70%.[8,9] This variability highlights the need for improved patient selection methods and enhanced understanding of factors influencing surgical success.

Genetic factors are increasingly recognized as key contributors to epilepsy susceptibility, severity, and treatment resistance. Mutations affecting ion channel function, synaptic transmission, cortical development, and neuroinflammation have been implicated in a broad spectrum of epilepsy syndromes.[10,11] These findings have opened the door to the use of genetic data in clinical decision-making, particularly in the context of presurgical evaluation.[12,13] The National Society of Genetic Counselors in the USA recommends genetic testing for all patients with unexplained epilepsy.[14] While testing is commonly prioritized for early-onset, syndromic, or developmental epilepsies, its role is expanding in focal DRE as well. For example, recent studies demonstrate that genetic diagnoses can directly influence management and outcomes, leading to treatment modifications in up to 45% of children with severe epilepsy – with around 60% of those receiving gene-directed therapies achieving seizure freedom – and informing clinical decisions in adults, where diagnostic yields range from 23% to 50%.[15,16] Specific findings now inform individualized approaches, such as avoiding sodium channel blockers in *SCN1A*-related epilepsy or using the ketogenic diet in *SLC2A1* cases, while emerging biomarkers like circulating microRNAs and protein profiles help distinguish drug-resistant from drug-sensitive epilepsy.[17–20] Despite growing interest, the incorporation of genetic biomarkers into routine epilepsy care remains limited in many LMICs because of infrastructure, cost, and workforce constraints.

This review summarizes the current evidence on the role of genetic biomarkers in DRE management, focusing on presurgical evaluation and treatment outcomes. Special attention is given to the challenges and opportunities in LMICs, where cost-effective, personalized approaches to epilepsy care are particularly critical. Through synthesis of the current literature, this review provides an overview of the genetic influences on DRE and explores how genetic testing can complement existing diagnostic strategies.

Objectives

The objective of this review is to explore how genetic testing can support presurgical evaluation in DRE, particularly in LMICs. It addresses the clinical impact, economic considerations, and feasibility of applying genetic screening to guide treatment decisions in resource-limited settings.

Materials and methods

A literature search was conducted across PubMed, Scopus, Web of Science, and Google Scholar. No restrictions were placed on publication date. Search terms included "drug-resistant epilepsy," "genetic testing," "epilepsy surgery," "presurgical evaluation," "clinical utility," "diagnostic yield," "cost-effectiveness," and "low- and middle-income countries."

The inclusion criteria comprised full-text, peer-reviewed articles in English describing the role of genetic testing in DRE, particularly in the context of presurgical decision-making and treatment planning. Studies involving both pediatric and adult populations were considered. Case reports, cohort studies, reviews, and expert consensus documents were included if they discussed the diagnostic yield, clinical utility, or cost-related implications of genetic testing. Editorials, commentaries, abstract-only records, and studies focusing exclusively on syndromic epilepsies without implications for surgical evaluation were excluded.

Due to the absence of studies conducted specifically in LMICs, this review incorporated findings from high-income countries (HICs), where most genetic testing protocols and outcome data are currently available. These studies were selected for their potential applicability to LMIC settings, especially regarding implementation strategies, diagnostic pathways, and resource prioritization. A systematic review was not feasible because of the limited number of regionally relevant studies and the heterogeneity of study designs and outcome reporting.

Genetic contributors to drug-resistant epilepsy

Genetic variants underlie many cases of severe and focal epilepsy. Ion channel variants, particularly in sodium channels (*SCN1A, SCN2A, SCN8A*) and potassium channels (*KCNQ2/3*), represent the most extensively studied genetic determinants of treatment response. De novo or inherited mutations in these genes cause early-onset epileptic encephalopathies and Dravet syndrome, which demonstrate inherent pharmacoresistance.[21] Recent genome-wide association studies have revealed that the *SCN1A* rs2298771 polymorphism significantly increases the risk of antiepileptic drugs (AED) resistance, particularly in South Asian populations.[22]

Synaptic genes (*STXBP1, PCDH19, CDKL5*) and metabolic genes (*SLC2A1, ALDH7A1*) also drive DRE through developmental encephalopathies or neurometabolic syndromes.[23–27] Drug transporter genetics have emerged as critical determinants of treatment response, with the *ABCB1* gene encoding P-glycoprotein showing polymorphisms (particularly *C3435T*) that significantly affect carbamazepine and lamotrigine transport across the blood–brain barrier (BBB).[28–30] Additionally, cytochrome P450 variants (*CYP2C9, CYP2C19*) influence AED metabolism and contribute to treatment resistance through altered drug clearance.[17]

Mutations in the mTOR pathway genes (*TSC1/2, DEPDC5, NPRL2/3, PTEN*) are linked to focal cortical dysplasia and tuberous sclerosis, often yielding localized lesions amenable to resection.[12,31] For example, *DEPDC5* mutations frequently co-occur with radiologically occult focal cortical dysplasia and predict the presence of surgically treatable malformations. In contrast, germline mutations in channelopathy genes (*SCN1A, CNTNAP2, STXBP1*) typically indicate diffuse network pathology and poor surgical outcomes.[13] Collectively, pathogenic variants in these genes can drive pharmacoresistance and inform the identification of epileptogenic zones.

Recent advances have increasingly highlighted the critical role of epigenetic modifications in the development and progression of DRE.[29,32] Specifically, DNA methylation patterns exhibit disease-specific alterations in hippocampal tissue obtained from patients with drug-resistant temporal lobe epilepsy.[33] These epigenetic changes influence the expression of genes encoding ion channels and neurotransmitter receptors, thereby contributing to the dysregulation of neuronal excitability and synaptic transmission characteristic of refractory seizures. In parallel, microRNA profiling has uncovered a set of circulating small noncoding RNAs, including miR-134-5p, miR-122-5p, and miR-132-3p, that reliably differentiate drug-resistant from drug-sensitive epilepsy patients.[19] These microRNAs are implicated in key neurobiological processes such as synaptic plasticity, neuroinflammation, and neuronal survival, all of which play pivotal roles in epileptogenesis and pharmacoresistance. The identification of these epigenetic markers offers promising avenues for developing minimally invasive diagnostic tools and novel therapeutic targets tailored to overcome drug resistance in epilepsy.

The growth of exome/genome sequencing has expanded the list of epilepsy genes; however, practical testing often focuses on established candidates. The ClinGen Epilepsy Gene Curation panel classifies approx. 30 genes as definitive causes of epilepsy (*SCN1A, KCNQ2, TSC1/2, DEPDC5, SCN2A, SLC2A1*), with many others supported by moderate evidence.[34] In pediatric DRE cohorts, diagnostic yields of targeted epilepsy gene panels range from around 12% to 29%, rising to 39–50% in infants with developmental encephalopathies. Whole-exome sequencing (WES) achieves even higher yields – up to 50% in early-onset, syndromic cases.[35] Notably, yields are higher when testing is performed early; 1 study found that testing at epilepsy onset halved the cost of the diagnostic workup by avoiding redundant tests.[36] Thus, for selected patients (early-onset, multifocal features, family history), genetic testing is highly informative (Table 1).

Genetic testing in the presurgical evaluation of DRE

Presurgical evaluation in DRE traditionally relies on clinical assessment, neuroimaging, and electroencephalography (EEG) to pinpoint the epileptogenic zone. However, even with these tools, surgical outcomes remain suboptimal in many patients, with seizures often persisting after surgery.[21,64,65] The addition of genetic data to this approach is an emerging paradigm. A novel approach to improve these outcomes is the integration of genetic testing into

Table 1. Key genetic biomarkers in drug-resistant epilepsy and their clinical implications

Gene/variant	Epilepsy phenotype	Testing availability	Surgical candidacy	Targeted treatment	References
SCN1A	Dravet syndrome	HIC: 95%, LMIC: 15–25%	poor (diffuse pathology)	avoid Na^+ blockers, stiripentol, clobazam	11,37–41
SCN2A	early infantile epileptic encephalopathy	HIC: 90%, LMIC: 10–20%	variable	Na^+ blockers (LOF), avoid in GOF	42,43
KCNQ2	benign familial neonatal seizures/EEIE	HIC: 85%, LMIC: 8–15%	poor to moderate	retigabine, phenytoin	34,44
TSC1/2	tuberous sclerosis complex	HIC: 95%, LMIC: 30–40%	good (focal lesions)	mTOR inhibitors, targeted resection	13,31,45–47
DEPDC5	focal cortical dysplasia	HIC: 80%, LMIC: 5–10%	excellent (focal lesions)	surgical resection	48–51
NPRL2/3	focal cortical dysplasia	HIC: 70%, LMIC: <5%	excellent	surgical resection	52,53
STXBP1	early infantile epileptic encephalopathy	HIC: 85%, LMIC: 10–15%	poor (diffuse pathology)	supportive care, avoid surgery	54,55
PCDH19	*PCDH19*-related epilepsy	HIC: 60%, LMIC: <5%	poor to moderate	hormone-based considerations	56
SLC2A1	GLUT1 deficiency	HIC: 90%, LMIC: 15–25%	variable	ketogenic diet	57–60
ALDH7A1	pyridoxine-dependent epilepsy	HIC: 85%, LMIC: 20–30%	variable	pyridoxine/pyridoxal phosphate	25,61–63

HIC – high-income countries; LMIC – low- and middle-income countries; LOF – loss of function; GOF – gain of function; EEIE – early epileptic and infantile encephalopathy; ID – intellectual disability.

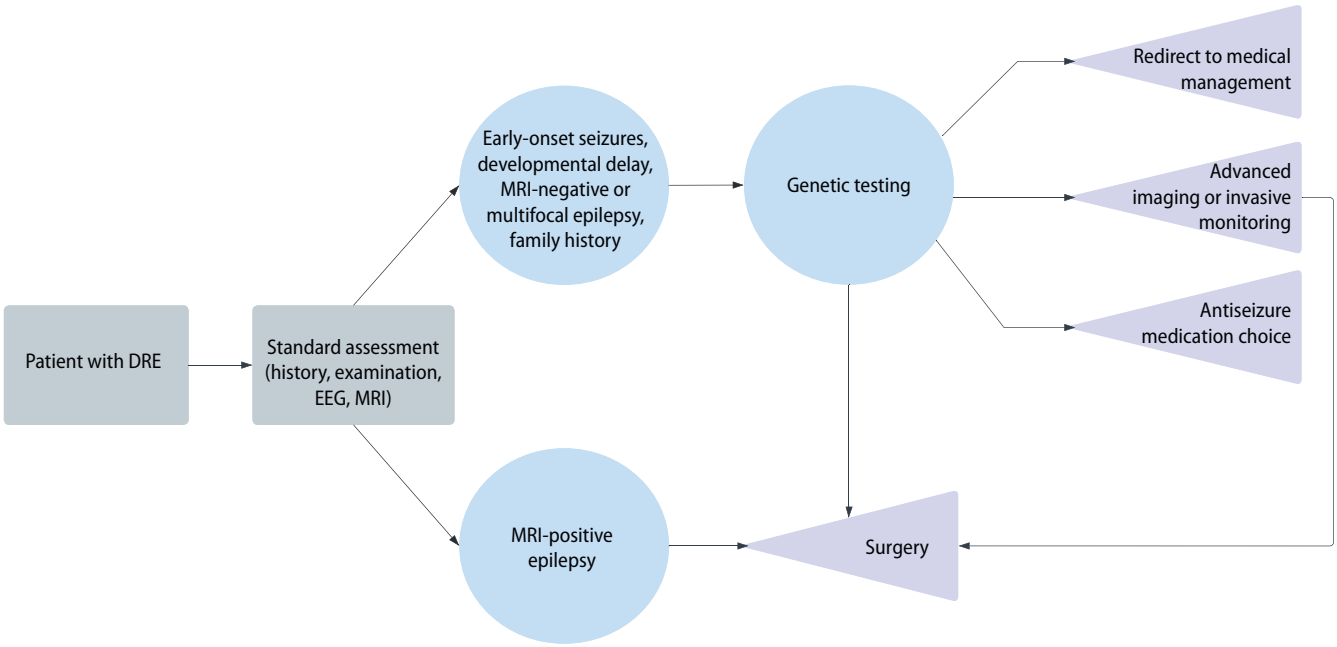

Fig. 1. Flowchart for integrating genetic testing into presurgical evaluation of drug-resistant epilepsy (DRE)

MRI – magnetic resonance imaging.

the presurgical workup (Fig. 1). Growing expert consensus and recent studies suggest that genetic testing should be considered in patients with focal epilepsy of unknown etiology – particularly those with early-onset seizures, developmental delay, or magnetic resonance imaging (MRI)-negative findings – where genetic etiologies may influence both surgical decision-making and outcomes.[66,67] Large studies underscore this shift: in a 5-year single-center cohort of 125 pediatric DRE surgical referrals, 69% underwent genetic testing, and 21% had a pathogenic variant identified.[68] The identified genes included *NPRL3, TSC2, KCNH1, CHRNA4, SPTAN1, DEPDC5, SCN2A, ARX, SCN1A, DLG4*, and *STXBP1*. Notably, in 3 cases, the genetic diagnosis directly altered management: those patients did not proceed to surgery because the findings indicated that a surgical cure was unlikely. The authors

emphasized that obtaining a molecular diagnosis can significantly impact treatment decisions and suggested that genetic testing should be integrated into the routine assessment process for pediatric patients with DRE during presurgical evaluation.

Other centers have reported similar findings. In 1 cohort, routine WES and microarray analysis of 49 children undergoing epilepsy surgery identified variants in 45% (21 of 49) of the children.[13] *TSC1/2, DEPDC5, MECP2*, and *KCNQ2*, among others, were identified by both exome and chromosomal microarray analyses. Although the genetic findings did not change whether surgery was performed, they guided postoperative management: patients with mutations were counseled more conservatively, retaining antiseizure medications for longer. Notably, among the few patients with poor surgical outcomes (International League Against Epilepsy (ILAE) class IV–V), a majority had underlying genetic variants or multifocal syndromes. This suggests that detecting a genetic cause can inform prognosis, guide patient counseling, and aid postoperative planning, even if it does not alter the immediate surgical plan.

Further evidence comes from smaller series. In a focused study of 9 children evaluated for focal epilepsy surgery, targeted epilepsy gene testing led to a major change in management in most cases: 7 of the 9 patients had their surgical workups halted entirely. All 7 had early-onset seizures with developmental delay, and each harbored a pathogenic germline variant in a known epilepsy gene.[67] The investigators recommend that genetic assessment should be incorporated as a mandatory component of presurgical evaluation, particularly in patients with negative MRI findings or ambiguous lesions. This recommendation aligns with the perspective that genetic diagnoses may contraindicate invasive monitoring or resection procedures, as these conditions typically represent diffuse neurological disease not amenable to focal surgical cure.[69]

Genetic stratification can improve presurgical decision-making in several ways. First, it complements imaging and EEG by revealing the underlying mechanism of epilepsy. For example, discovering an mTOR pathway mutation (*DEPDC5, TSC1/2*) in an MRI-negative patient may prompt the use of invasive EEG (stereo-EEG) to detect subtle focal cortical dysplasia.[31] Conversely, identifying a "generalized" genetic epilepsy (e.g., *SCN1A, STXBP1*) suggests poor localization, leading clinicians to consider medical or neuromodulation therapies instead of surgery.[69] Second, genetic findings can directly affect pharmacological management. A well-known example is that *SCN1A* mutations contraindicate sodium-channel blockers,[37] whereas *SCN2A/SCN8A* mutation (when loss-of-function) may benefit from them.[42,43] Similarly, detecting *SLC2A1* (*GLUT1* deficiency) guides dietary therapy (ketogenic diet),[57] and *ALDH7A1* identifies patients treatable with pyridoxine.[61] Awareness of such gene-specific therapies can improve seizure control before and after surgery. Finally, knowledge of genetic drug transport or metabolism variants (*ABCB1* transporter polymorphisms) may explain resistance and indicate alternative drug choices.[70]

Mounting clinical evidence reinforces the value of these approaches. Multiple studies have found that obtaining a genetic diagnosis leads to changes in medical management in nearly half of DRE patients and can improve seizure outcomes in a substantial proportion of cases.[71] Moreover, certain genetic findings correlate with surgical prognosis: e.g., genetic disorders that cause localized structural abnormalities (such as mTOR-pathway focal cortical dysplasias or mesial temporal sclerosis) are associated with more favorable post-surgical outcomes, whereas patients with genetic generalized epilepsies tend to have poorer seizure control after resection. Taken together, this evidence strongly supports the routine incorporation of genetic testing into presurgical evaluation for DRE – particularly in pediatric patients and those without evident lesions on MRI – to ensure that each patient's treatment plan is informed by the underlying cause of their epilepsy.

Integrating genetic testing into clinical pathways

Given the benefits genetic information can provide, many epilepsy centers are now integrating genetic testing into their standard clinical pathways for DRE evaluations. In current practice – especially for pediatric DRE cases – a streamlined presurgical workflow has emerged that incorporates genetic testing early in the diagnostic process.[36] Following standard clinical, neuroimaging, and electrophysiological assessments, genetic screening is increasingly utilized early in the diagnostic process for individuals meeting high-priority criteria, such as seizure onset before the age of 2, positive family history, failure of 2 or more anti-seizure medications, or normal MRI findings. In many centers, targeted epilepsy gene panels or WES are ordered concurrently with traditional diagnostic evaluations.[68,72] Pathogenic variants identified using these methods may redirect clinical management by precluding invasive monitoring or surgery in cases of diffuse genetic epilepsy or by supporting surgical candidacy in patients with focal lesions. Even when the results are negative or of uncertain significance, they may guide the consideration of additional testing modalities, including assessments for copy number variants or mosaicism. Genetic findings are interpreted in conjunction with neuroimaging and EEG data within multidisciplinary epilepsy teams, which increasingly include expertise in clinical genetics to support integrated decision-making.[73]

Early experiences with this integrated approach have shown clear advantages. Recent case series highlight that performing genetic testing sooner rather than later in the evaluation can streamline care and reduce unnecessary procedures. For example, 1 study found that when genetic testing was conducted within the 1st year of a child's epilepsy (early in the disease course), patients

required significantly fewer metabolic tests and invasive investigations to reach a diagnosis, compared to those who were tested only after multiple years of DRE.[36] In that 5-year series, the overall diagnostic yield of genetic testing was about 12% (28 of 226 patients received a genetic diagnosis). Notably, the subgroup of patients who had early genetic testing accounted for 8 of those diagnoses and avoided numerous additional tests, resulting in lower overall healthcare costs. This suggests that even with a moderate diagnostic yield, early genetic testing can act as a cost-effective triage tool – streamlining the workup by ruling in or out certain etiologies and thereby preventing more costly or invasive diagnostic procedures. As the cost of sequencing continues to fall (WES can now often be performed for under $1,000) and bioinformatics pipelines improve, incorporating genetics early has become increasingly feasible in routine practice.[74] Furthermore, economic models from other areas of medicine have shown that first-tier genomic testing can be cost-effective under standard willingness-to-pay thresholds, supporting its use as an upfront diagnostic strategy in appropriate cases.[75]

In resource-limited environments, the cost-effectiveness of genetic screening must account for local factors. The upfront cost of a gene panel or WES may appear high, but it must be weighed against the costs of failed surgeries, prolonged hospital evaluations, and serial medication trials. For example, epilepsy surgery costs in LMICs range from $500 to $8,000 depending on the country, yet these figures often reflect direct costs without accounting for presurgical evaluations, hospital stays, or postoperative care. Additionally, many clinicians in LMICs are hesitant to refer patients for epilepsy surgery due to limited training and uncertainty about outcomes, which means resources for surgical evaluation may already be underutilized.[76] A genetic diagnosis that prevents an unsuccessful surgery can save much more than the test itself. One health economic analyses in pediatric epilepsy (largely from high-income settings) indicated that identifying genetic causes early reduces the diagnostic odyssey and improves long-term outcomes, yielding overall cost.[77] Such conclusions likely hold in LMICs, where inefficiencies carry even greater relative burdens.

Overall, the inclusion of genetic testing in standard clinical pathways for DRE allows clinicians to make more informed decisions earlier. This leads to more efficient use of diagnostic resources, prevents patients from undergoing unnecessary invasive procedures or ineffective surgeries, and ensures that each patient's care plan (medical or surgical) is optimized based on genetic findings.

Challenges and opportunities in LMIC implementation

Despite the promising benefits of genetic testing, implementing these advances in LMICs (Fig. 2) presents unique challenges. Healthcare systems in LMICs often face resource constraints that can hinder the adoption of routine genetic testing in epilepsy care. Infrastructure limitations are a major hurdle: advanced neuroimaging (high-resolution MRI) and long-term EEG monitoring are scarce in many regions, and local molecular genetics laboratories are often lacking. A recent global survey of 1,568 providers across 127 countries illustrates the extent of the gap. It found that advanced genetic testing (such as WES or whole-genome sequencing (WGS)) was available to 90.4% of epilepsy care providers in HICs, but only 16.3% of providers in LMICs had access to these technologies.[78] In Africa, only 12% reported access to gene panels, whereas in Latin America, the figure was 26% (compared to 63% in Europe). Even where tests exist, most LMICs lack public funding; 37.5% of providers in low-income settings reported that patients must self-fund genetic tests, compared to only 8.6% in high-income settings.[78] The high

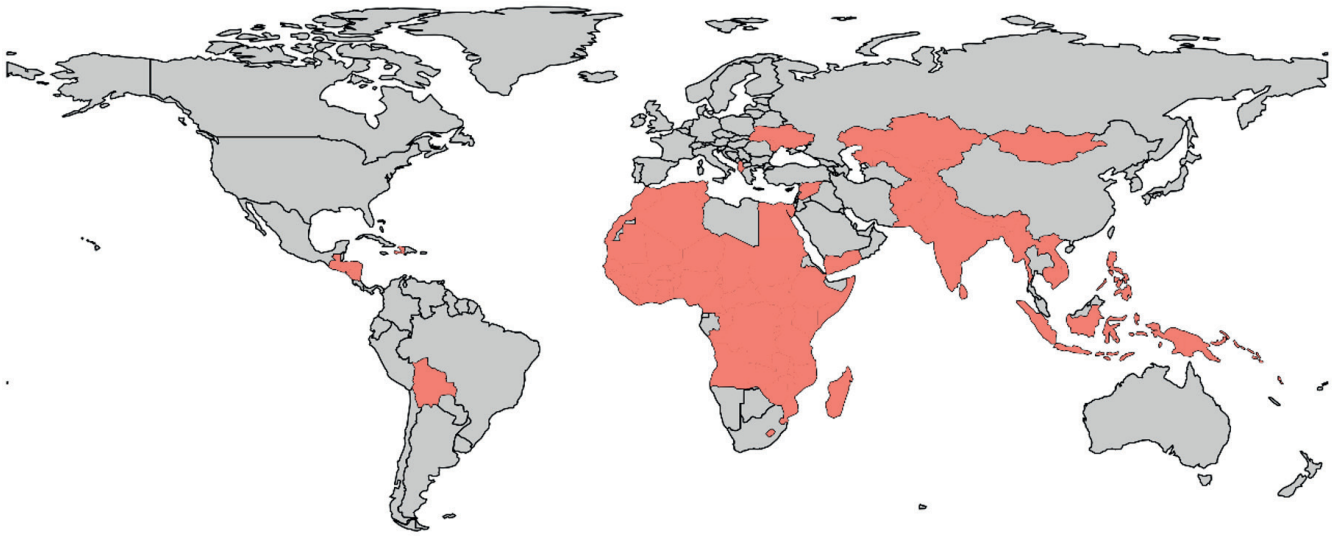

Fig. 2. Low- and middle-income countries (LMICs)

cost of sequencing and limited laboratory infrastructure are clearly key drivers of these disparities.

Another challenge is that genetic data interpretation can be more difficult in LMIC populations. Global reference databases for genetic variants are heavily skewed toward individuals of European ancestry, which means that variants found in patients from other ethnic backgrounds (common in Africa, Asia, and Latin America) are often classified as "variants of uncertain significance" due to a lack of comparative data.[79] This uncertainty can limit the clinical actionability of genetic results in non-European populations. Additionally, awareness and training in epilepsy genetics are limited among healthcare providers in many LMICs. Clinicians may not be familiar with when or how to order genetic tests, and there may be cultural or social stigma around genetic disorders that makes families hesitant to pursue testing.[80–82] These factors all contribute to a slower uptake of genetic testing in routine practice.

Despite these barriers, there are still promising developments. International collaborations (ILAE task forces and non-governmental organizations (NGOs)) have established centers of excellence for epilepsy genetics in Africa, Asia, and Latin America (https://www.ilae.org). Telemedicine and cloud-based analysis can compensate for shortages in local expertise.[83] Some countries are piloting targeted epilepsy panels using portable sequencers or in partnerships with foreign laboratories.[84] Moreover, LMICs can adopt a phased, high-impact approach and prioritize genetic testing for high-yield cases. For example, testing could be reserved for early-onset DRE, patients with a suggestive family history, and cases in which surgery is being considered. In such targeted cohorts, diagnostic yields of genetic testing are often in the 20–40% range, which makes testing cost-effective given the potential to change management. For example, a recent study in Nigeria performed exome sequencing on 22 children with unexplained DRE and achieved a 27% diagnostic rate.[85] These 6 diagnoses (genes *BPTF, NAA15, SCN1A, TUBA1A, CACNA1A*) not only explain epilepsy but also open potential therapeutic avenues (e.g., channel blockers for *CACNA1A*-related seizures). This kind of focused testing demonstrates that even in LMIC settings, genetic diagnostics can yield clinically meaningful results when applied to high-yield scenarios.

The economic rationale for genetics in LMICs is compelling. Identifying a mutation that predicts surgical failure prevents unnecessary expenditure on invasive evaluations and surgeries for that patient, thereby conserving limited resources. Conversely, identifying resectable genetic epilepsy can help prioritize patients requiring limited surgery slots. As sequencing becomes increasingly accessible, LMICs will benefit from precision methodologies. Key requirements for success in LMICs include capacity-building (training neurologists and geneticists, and educating general clinicians about genetics), securing funding support (through government programs or international donors) to subsidize testing costs, and adapting clinical guidelines to local realities (e.g., creating simplified genetic testing protocols that can be followed in low-resource hospitals). Notably, the WHO's Global Action Plan on epilepsy (2022–2031) explicitly calls for timely diagnosis and care across all regions,[4] implying that access to genetic tools should be part of this global strategy.

Future directions

Looking ahead, the role of genetic testing in managing DRE is set to expand, and several important directions for future development are evident. One critical need is for large multicenter studies that can validate specific genetic biomarkers as predictors of surgical outcomes and inform evidence-based guidelines. Thus far, most studies linking genetics to epilepsy surgery outcomes have been relatively small, focused primarily on pediatric cases, and often single-center. To generalize these findings, future research should include more diverse patient populations (including adults and various ethnic groups) and incorporate data from LMICs, where genetic contributions to epilepsy might differ. These larger studies would enable the development of standardized screening algorithms – tools that help clinicians decide which DRE patients should undergo genetic testing and how to interpret the results in the context of surgical decision-making. Additionally, further cost-effectiveness analyses are needed, especially those tailored to low-resource healthcare systems, to guide policymakers on how to implement genetic testing in a sustainable way.

Emerging technologies will likely accelerate progress in this field. Portable and point-of-care genetic sequencing devices are being developed, which could allow clinics (even in remote areas) to perform genetic tests without relying on distant laboratories. Similarly, advances in bioinformatics and artificial intelligence (AI)-driven variant interpretation are expected to make it faster and easier to distinguish pathogenic mutations from benign variants, addressing one of the current bottlenecks in genomic medicine. As these tools become more affordable and user-friendly, they will help bring genetic diagnostics into mainstream clinical use. At the same time, it will be important to address ethical, cultural, and privacy considerations surrounding genetic testing. Engaging communities to improve understanding of epilepsy genetics, tackling stigma associated with genetic conditions, ensuring informed consent, and establishing clear guidelines for data privacy and sharing are all essential steps as broader implementation of genomic medicine in epilepsy progresses.

While we anticipate these future developments, there are steps that epilepsy programs can take now to begin harnessing genetics. We recommend that multidisciplinary epilepsy centers develop and adopt protocols for selective genetic screening as part of the DRE evaluation. One model is reflexive testing: for instance, automatically sending a blood sample for WES as soon as a patient is diagnosed

with DRE (after failing 2 appropriate medications), rather than waiting for multiple failed treatments or inconclusive tests. Another approach is to use targeted gene panels for specific clinical scenarios (e.g., a panel of known infantile epilepsy genes for a child with seizures starting in the 1st year of life).

Moreover, increasing collaboration between hospitals in LMICs and international research consortia can provide interim solutions. Initiatives such as H3Africa (https://h3africa.org) and Central Asian & Transcaucasian Genomics (https://www.cat-genomics.com) are already connecting clinicians in developing regions with resources and expertise in genomics. By partnering with such consortia, a center in a low-resource setting can access genetic sequencing and interpretation services for its patients, even as it works on building local capacity. Importantly, whenever a genetic diagnosis is made, it must be accompanied by appropriate genetic counseling and integrated clinical follow-up. The ultimate goal is not just to identify genes for the sake of knowledge, but to translate that knowledge into better patient care – whether it means altering medications, recommending dietary therapy, advising against unnecessary surgery, or screening family members at risk.

Limitations of the study

This review has several limitations. First, the synthesis is based primarily on studies conducted in high-income settings, which may not be generalizable to LMIC contexts due to differences in healthcare infrastructure, genetic ancestry, and treatment availability. The scarcity of large, population-based studies from LMICs limited our ability to draw region-specific conclusions or assess implementation feasibility in these settings. Many of the included reports were case series or single-center studies with small sample sizes, introducing potential selection and publication biases. Cost-effectiveness data were extrapolated from HICs and may not reflect local economic conditions or healthcare priorities in LMICs. Furthermore, the review did not include a systematic assessment of study quality or risk of bias due to the heterogeneity and limited number of eligible studies. The absence of standardized outcome measures and inconsistent reporting of genetic variants also hindered comparisons across studies. Finally, ethical and sociocultural considerations – such as stigma, informed consent, and data sharing – remain underexplored in the existing literature, particularly in LMIC contexts, and merit dedicated investigation in future research.

Conclusions

Genetic biomarker screening holds significant potential to improve the management of DRE, especially when integrated into presurgical evaluation. The evidence synthesized in this review indicates that incorporating genetic testing into standard workups can refine patient selection for epilepsy surgery and guide more personalized treatment strategies. By identifying underlying genetic etiologies, clinicians are better able to predict which patients are likely to benefit from surgical intervention and which may require alternative approaches, thereby avoiding futile surgeries and optimizing the use of limited resources. This approach is particularly impactful in LMICs, where healthcare resources are constrained and the consequences of unsuccessful interventions are considerable. As genetic testing becomes more accessible and cost-effective, its adoption in presurgical protocols across diverse settings could help close the epilepsy treatment gap. Integrating genetics into routine DRE care paves the way for precision medicine approaches that improve patient outcomes, reduce the burden of uncontrolled seizures, and ultimately enhance the quality of life for individuals with DRE worldwide.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Dina Kalinina https://orcid.org/0000-0003-3887-1170
Alimzhan Muxunov https://orcid.org/0009-0001-8860-0363
Zhassulan Utebekov https://orcid.org/0009-0006-6272-5033
Gaziz Kyrgyzbay https://orcid.org/0009-0003-1156-9374
Darkhan Kimadiev https://orcid.org/0000-0002-5490-8642
Guldana Zhumabaeva https://orcid.org/0000-0002-5511-8060
Joseph Almazan https://orcid.org/0000-0001-5148-6889
Antonio Sarria-Santamera https://orcid.org/0000-0001-5734-7468

References

1. Meyer AC, Dua T, Ma J, Saxena S, Birbeck G. Global disparities in the epilepsy treatment gap: A systematic review. *Bull World Health Organ*. 2010;88(4):260–266. doi:10.2471/BLT.09.064147
2. Feigin VL, Vos T, Nair BS, et al. Global, regional, and national burden of epilepsy, 1990–2021: A systematic analysis for the Global Burden of Disease Study 2021. *Lancet Public Health*. 2025;10(3):e203–e227. doi:10.1016/S2468-2667(24)00302-5
3. Beniczky S, Trinka E, Wirrell E, et al. Updated classification of epileptic seizures: Position paper of the International League Against Epilepsy. *Epilepsia*. 2025;66(6):1804–1823. doi:10.1111/epi.18338
4. World Health Organization (WHO). *Intersectoral Global Action Plan on Epilepsy and Other Neurological Disorders*. Geneva, Switzerland: World Health Organization (WHO); 2023. https://iris.who.int/bitstream/handle/10665/371495/9789240076624-eng.pdf?sequence=1.
5. Leonardi M, Ustun TB. The global burden of epilepsy. *Epilepsia*. 2002;43(Suppl 6):21–25. doi:10.1046/j.1528-1157.43.s.6.11.x
6. Sabzvari T, Aflahe Iqbal M, Ranganatha A, et al. A comprehensive review of recent trends in surgical approaches for epilepsy management. *Cureus*. 2024;16(10):e71715. doi:10.7759/cureus.71715
7. Maas L, Kellenaers J, Van Mastrigt G, et al. Societal costs and quality of life analysis in patients undergoing resective epilepsy surgery: A one-year follow-up. *Epilepsy Behav Rep*. 2023;24:100635. doi:10.1016/j.ebr.2023.100635
8. Englot DJ, Chang EF. Rates and predictors of seizure freedom in resective epilepsy surgery: An update. *Neurosurg Rev*. 2014;37(3):389–405. doi:10.1007/s10143-014-0527-9
9. Salem T, Chetty C, Chetty O. A review of the surgical procedures for the treatment of drug-resistant epilepsy and their seizure control outcomes. *Surg Sci*. 2023;14(8):533–549. doi:10.4236/ss.2023.148058

10. Chen T, Giri M, Xia Z, Subedi YN, Li Y. Genetic and epigenetic mechanisms of epilepsy: A review. *Neuropsychiatr Dis Treat*. 2017;13:1841–1859. doi:10.2147/NDT.S142032
11. Oyrer J, Maljevic S, Scheffer IE, Berkovic SF, Petrou S, Reid CA. Ion channels in genetic epilepsy: From genes and mechanisms to disease-targeted therapies. *Pharmacol Rev*. 2018;70(1):142–173. doi:10.1124/pr.117.014456
12. Sanders MWCB, Lemmens CMC, Jansen FE, et al. Implications of genetic diagnostics in epilepsy surgery candidates: A single-center cohort study. *Epilepsia Open*. 2019;4(4):609–617. doi:10.1002/epi4.12366
13. Becker L, Makridis KL, Abad-Perez AT, et al. The importance of routine genetic testing in pediatric epilepsy surgery. *Epilepsia Open*. 2024;9(2):800–807. doi:10.1002/epi4.12916
14. Smith L, Malinowski J, Ceulemans S, et al. Genetic testing and counseling for the unexplained epilepsies: An evidence-based practice guideline of the National Society of Genetic Counselors. *J Genet Couns*. 2023;32(2):266–280. doi:10.1002/jgc4.1646
15. Sasaki E, Millington P, Sazonova T, et al. Impact of rapid genomic testing on clinical outcomes of acutely unwell children presenting with severe epilepsy [published online as ahead of print on May 21, 2025]. *Eur J Hum Genet*. 2025. doi:10.1038/s41431-025-01870-5
16. Clayton LM, Vakrinou A, Balestrini S, Sisodiya SM. Monogenic epilepsies in adult epilepsy clinics and gene-driven approaches to treatment. *Curr Neurol Neurosci Rep*. 2025;25(1):35. doi:10.1007/s11910-025-01413-x
17. Weber YG, Nies AT, Schwab M, Lerche H. Genetic biomarkers in epilepsy. *Neurotherapeutics*. 2014;11(2):324–333. doi:10.1007/s13311-014-0262-5
18. Knowles JK, Helbig I, Metcalf CS, et al. Precision medicine for genetic epilepsy on the horizon: Recent advances, present challenges, and suggestions for continued progress. *Epilepsia*. 2022;63(10):2461–2475. doi:10.1111/epi.17332
19. Timechko EE, Lysova KD, Yakimov AM, et al. Circulating microRNAs as biomarkers of various forms of epilepsy. *Med Sci (Basel)*. 2025;13(1):7. doi:10.3390/medsci13010007
20. Ma M, Cheng Y, Hou X, et al. Serum biomarkers in patients with drug-resistant epilepsy: A proteomics-based analysis. *Front Neurol*. 2024;15:1383023. doi:10.3389/fneur.2024.1383023
21. Karakis I. Genetic testing in the presurgical evaluation of drug-resistant epilepsy: Bells and whistles or nuts and bolts? *Epilepsy Curr*. 2024;24(4):248–250. doi:10.1177/15357597241250161
22. Li M, Zhong R, Lu Y, Zhao Q, Li G, Lin W. Association between SCN1A rs2298771, SCN1A rs10188577, SCN2A rs17183814, and SCN2A rs2304016 polymorphisms and responsiveness to antiepileptic drugs: A meta-analysis. *Front Neurol*. 2021;11:591828. doi:10.3389/fneur.2020.591828
23. Bartolini E. Inherited developmental and epileptic encephalopathies. *Neurol Int*. 2021;13(4):555–568. doi:10.3390/neurolint13040055
24. Spagnoli C, Fusco C, Percesepe A, Leuzzi V, Pisani F. Genetic neonatal-onset epilepsies and developmental/epileptic encephalopathies with movement disorders: A systematic review. *Int J Mol Sci*. 2021;22(8):4202. doi:10.3390/ijms22084202
25. Böhm HO, Yazdani M, Sandås EM, et al. Global metabolomics discovers two novel biomarkers in pyridoxine-dependent epilepsy caused by ALDH7A1 deficiency. *Int J Mol Sci*. 2022;23(24):16061. doi:10.3390/ijms232416061
26. Spagnoli C, Fusco C, Pisani F. Pediatric-onset epilepsy and developmental epileptic encephalopathies followed by early-onset parkinsonism. *Int J Mol Sci*. 2023;24(4):3796. doi:10.3390/ijms24043796
27. Cavirani B, Spagnoli C, Caraffi SG, et al. Genetic epilepsies and developmental epileptic encephalopathies with early onset: A multicenter study. *Int J Mol Sci*. 2024;25(2):1248. doi:10.3390/ijms25021248
28. Shahid A, Hameed K, Zainab A, Zafar A, Abbas S. Advances in pharmacogenomics: Optimizing antiepileptic drug therapy for drug-resistant epilepsy. *Explor Neuroprotect Ther*. 2024;4:240–250. doi:10.37349/ent.2024.00080
29. Bazhanova ED, Kozlov AA, Litovchenko AV. Mechanisms of drug resistance in the pathogenesis of epilepsy: Role of neuroinflammation. A literature review. *Brain Sciences*. 2021;11(5):663. doi:10.3390/brainsci11050663
30. Sisodiya SM, Lin WR, Harding BN, Squier MV, Thom M. Drug resistance in epilepsy: Expression of drug resistance proteins in common causes of refractory epilepsy. *Brain*. 2002;125(1):22–31. doi:10.1093/brain/awf002
31. Oane I, Barborica A, Daneasa A, et al. Organization of the epileptogenic zone and signal analysis at seizure onset in patients with drug-resistant epilepsy due to focal cortical dysplasia with mTOR pathway gene mutations: An SEEG study. *Epilepsia Open*. 2023;8(4):1588–1595. doi:10.1002/epi4.12810
32. Fonseca-Barriendos D, Frías-Soria CL, Pérez-Pérez D, Gómez-López R, Borroto Escuela DO, Rocha L. Drug-resistant epilepsy: Drug target hypothesis and beyond the receptors. *Epilepsia Open*. 2022;7(Suppl 1):S23–S33. doi:10.1002/epi4.12539
33. Miller-Delaney SFC, Bryan K, Das S, et al. Differential DNA methylation profiles of coding and non-coding genes define hippocampal sclerosis in human temporal lobe epilepsy. *Brain*. 2015;138(3):616–631. doi:10.1093/brain/awu373
34. Helbig I, Riggs ER, Barry C, et al. The ClinGen Epilepsy Gene Curation Expert Panel: Bridging the divide between clinical domain knowledge and formal gene curation criteria. *Hum Mutat*. 2018;39(11):1476–1484. doi:10.1002/humu.23632
35. Habela CW, Schatz K, Kelley SA. Genetic testing in epilepsy: Improving outcomes and informing gaps in research. *Epilepsy Curr*. 2025;25(3):153–157. doi:10.1177/15357597241232881
36. Swartwood SM, Morales A, Hatchell KE, et al. Early genetic testing in pediatric epilepsy: Diagnostic and cost implications. *Epilepsia Open*. 2024;9(1):439–444. doi:10.1002/epi4.12878
37. Catterall WA. Dravet syndrome: A sodium channel interneuronopathy. *Curr Opin Psychol*. 2018;2:42–50. doi:10.1016/j.cophys.2017.12.007
38. Vezyroglou A, Varadkar S, Bast T, et al. Focal epilepsy in *SCN1A*-mutation carrying patients: Is there a role for epilepsy surgery? *Dev Med Child Neurol*. 2020;62(11):1331–1335. doi:10.1111/dmcn.14588
39. Fan HC, Yang MT, Lin LC, Chiang KL, Chen CM. Clinical and genetic features of Dravet syndrome: A prime example of the role of precision medicine in genetic epilepsy. *Int J Mol Sci*. 2023;25(1):31. doi:10.3390/ijms25010031
40. Wirrell EC, Nabbout R. Recent advances in the drug treatment of Dravet syndrome. *CNS Drugs*. 2019;33(9):867–881. doi:10.1007/s40263-019-00666-8
41. Connolly MB. Dravet syndrome: Diagnosis and long-term course. *Can J Neurol Sci*. 2016;43(Suppl 3):S3–S8. doi:10.1017/cjn.2016.243
42. Ademuwagun IA, Rotimi SO, Syrbe S, Ajamma YU, Adebiyi E. Voltage gated sodium channel genes in epilepsy: Mutations, functional studies, and treatment dimensions. *Front Neurol*. 2021;12:600050. doi:10.3389/fneur.2021.600050
43. Berecki G, Howell KB, Heighway J, et al. Functional correlates of clinical phenotype and severity in recurrent *SCN2A* variants. *Commun Biol*. 2022;5(1):515. doi:10.1038/s42003-022-03454-1
44. Millichap JJ, Park KL, Tsuchida T, et al. *KCNQ2* encephalopathy: Features, mutational hot spots, and ezogabine treatment of 11 patients. *Neurol Genet*. 2016;2(5):e96. doi:10.1212/NXG.0000000000000096
45. Curatolo P, Moavero R, De Vries PJ. Neurological and neuropsychiatric aspects of tuberous sclerosis complex. *Lancet Neurol*. 2015;14(7):733–745. doi:10.1016/S1474-4422(15)00069-1
46. French JA, Lawson JA, Yapici Z, et al. Adjunctive everolimus therapy for treatment-resistant focal-onset seizures associated with tuberous sclerosis (EXIST-3): A phase 3, randomised, double-blind, placebo-controlled study. *Lancet*. 2016;388(10056):2153–2163. doi:10.1016/S0140-6736(16)31419-2
47. Krueger DA, Northrup H, Roberds S, et al. Tuberous sclerosis complex surveillance and management: Recommendations of the 2012 International Tuberous Sclerosis Complex Consensus Conference. *Pediatr Neurol*. 2013;49(4):255–265. doi:10.1016/j.pediatrneurol.2013.08.002
48. Baulac S, Ishida S, Marsan E, et al. Familial focal epilepsy with focal cortical dysplasia due to *DEPDC 5* mutations. *Ann Neurol*. 2015;77(4):675–683. doi:10.1002/ana.24368
49. Ribierre T, Deleuze C, Bacq A, et al. Second-hit mosaic mutation in mTORC1 repressor DEPDC5 causes focal cortical dysplasia-associated epilepsy. *J Clin Invest*. 2018;128(6):2452–2458. doi:10.1172/JCI99384
50. D'Gama AM, Geng Y, Couto JA, et al. Mammalian target of rapamycin pathway mutations cause hemimegalencephaly and focal cortical dysplasia. *Ann Neurol*. 2015;77(4):720–725. doi:10.1002/ana.24357
51. Samanta D. DEPDC5-related epilepsy: A comprehensive review. *Epilepsy Behav*. 2022;130:108678. doi:10.1016/j.yebeh.2022.108678
52. Baldassari S, Licchetta L, Tinuper P, Bisulli F, Pippucci T. GATOR1 complex: The common genetic actor in focal epilepsies. *J Med Genet*. 2016;53(8):503–510. doi:10.1136/jmedgenet-2016-103883

53. Koboldt DC, Steinberg KM, Larson DE, Wilson RK, Mardis ER. The next-generation sequencing revolution and its impact on genomics. *Cell*. 2013;155(1):27–38. doi:10.1016/j.cell.2013.09.006
54. Saitsu H, Kato M, Mizuguchi T, et al. De novo mutations in the gene encoding STXBP1 (MUNC18-1) cause early infantile epileptic encephalopathy. *Nat Genet*. 2008;40(6):782–788. doi:10.1038/ng.150
55. Stamberger H, Nikanorova M, Willemsen MH, et al. *STXBP1* encephalopathy: A neurodevelopmental disorder including epilepsy. *Neurology*. 2016;86(10):954–962. doi:10.1212/WNL.0000000000002457
56. Trivisano M, Pietrafusa N, Ciommo VD, et al. PCDH19-related epilepsy and Dravet syndrome: Face-off between two early-onset epilepsies with fever sensitivity. *Epilepsy Res*. 2016;125:32–36. doi:10.1016/j.eplepsyres.2016.05.015
57. Daci A, Bozalija A, Jashari F, Krasniqi S. Individualizing treatment approaches for epileptic patients with glucose transporter type 1 (GLUT-1) deficiency. *Int J Mol Sci*. 2018;19(1):122. doi:10.3390/ijms19010122
58. Pearson TS, Akman C, Hinton VJ, Engelstad K, De Vivo DC. Phenotypic spectrum of glucose transporter type 1 deficiency syndrome (Glut1 DS). *Curr Neurol Neurosci Rep*. 2013;13(4):342. doi:10.1007/s11910-013-0342-7
59. Klepper J, Akman C, Armeno M, et al. Glut1 deficiency syndrome (Glut1DS): State of the art in 2020 and recommendations of the international Glut1DS study group. *Epilepsia Open*. 2020;5(3):354–365. doi:10.1002/epi4.12414
60. Leen WG, Klepper J, Verbeek MM, et al. Glucose transporter-1 deficiency syndrome: The expanding clinical and genetic spectrum of a treatable disorder. *Brain*. 2010;133(3):655–670. doi:10.1093/brain/awp336
61. Gospe SM. Pyridoxine-dependent epilepsy: ALDH7A1. In: Adam MP, Feldman J, Mirzaa GM, Pagon RA, Wallace SE, Amemiya A, eds. *GeneReviews®*. Seattle, USA: University of Washington; 1993. http://www.ncbi.nlm.nih.gov/books/NBK1486. Accessed August 22, 2025.
62. Mills PB, Struys E, Jakobs C, et al. Mutations in antiquitin in individuals with pyridoxine-dependent seizures. *Nat Med*. 2006;12(3):307–309. doi:10.1038/nm1366
63. Van Karnebeek CDM, Tiebout SA, Niermeijer J, et al. Pyridoxine-dependent epilepsy: An expanding clinical spectrum. *Pediatr Neurol*. 2016;59:6–12. doi:10.1016/j.pediatrneurol.2015.12.013
64. Englot DJ, Nagarajan SS, Imber BS, et al. Epileptogenic zone localization using magnetoencephalography predicts seizure freedom in epilepsy surgery. *Epilepsia*. 2015;56(6):949–958. doi:10.1111/epi.13002
65. Goldenholz DM, Jow A, Khan OI, et al. Preoperative prediction of temporal lobe epilepsy surgery outcome. *Epilepsy Res*. 2016;127:331–338. doi:10.1016/j.eplepsyres.2016.09.015
66. Ottman R, Hirose S, Jain S, et al. Genetic testing in the epilepsies: Report of the ILAE Genetics Commission. *Epilepsia*. 2010;51(4):655–670. doi:10.1111/j.1528-1167.2009.02429.x
67. Garcia-Uzquiano R, Barcia G, Losito E, et al. Genetic testing, another important tool in presurgical evaluation of focal epilepsies in childhood. *Epilepsia Open*. 2024;9(4):1589–1596. doi:10.1002/epi4.12964
68. Alsubhi S, Berrahmoune S, Dudley RWR, et al. Utility of genetic testing in the pre-surgical evaluation of children with drug-resistant epilepsy. *J Neurol*. 2024;271(5):2503–2508. doi:10.1007/s00415-023-12174-3
69. Straka B, Splitkova B, Vlckova M, et al. Genetic testing in children enrolled in epilepsy surgery program: A real-life study. *Eur J Paediatr Neurol*. 2023;47:80–87. doi:10.1016/j.ejpn.2023.09.009
70. Siddiqui A, Kerb R, Weale ME, et al. Association of multidrug resistance in epilepsy with a polymorphism in the drug-transporter gene *ABCB1*. *N Engl J Med*. 2003;348(15):1442–1448. doi:10.1056/NEJMoa021986
71. McKnight D, Morales A, Hatchell KE, et al. Genetic testing to inform epilepsy treatment management from an international study of clinical practice. *JAMA Neurol*. 2022;79(12):1267. doi:10.1001/jamaneurol.2022.3651
72. Moloney PB, Dugan P, Widdess-Walsh P, Devinsky O, Delanty N. Genomics in the presurgical epilepsy evaluation. *Epilepsy Res*. 2022;184:106951. doi:10.1016/j.eplepsyres.2022.106951
73. Vadlamudi L, Bennett CM, Tom M, et al. A multi-disciplinary team approach to genomic testing for drug-resistant epilepsy patients: The GENIE Study. *J Clin Med*. 2022;11(14):4238. doi:10.3390/jcm11144238
74. Schwarze K, Buchanan J, Taylor JC, Wordsworth S. Are whole-exome and whole-genome sequencing approaches cost-effective? A systematic review of the literature. *Genet Med*. 2018;20(10):1122–1130. doi:10.1038/gim.2017.247
75. Nurchis MC, Radio FC, Salmasi L, et al. Cost-effectiveness of whole-genome vs whole-exome sequencing among children with suspected genetic disorders. *JAMA Netw Open*. 2024;7(1):e2353514. doi:10.1001/jamanetworkopen.2023.53514
76. Watila MM, Xiao F, Keezer MR, et al. Epilepsy surgery in low- and middle-income countries: A scoping review. *Epilepsy Behav*. 2019;92:311–326. doi:10.1016/j.yebeh.2019.01.001
77. Howell KB, Eggers S, Dalziel K, et al. A population-based cost-effectiveness study of early genetic testing in severe epilepsies of infancy. *Epilepsia*. 2018;59(6):1177–1187. doi:10.1111/epi.14087
78. Reid-Bashir M, Esterhuizen A, Hildebrand M, et al. International study of epilepsy genetic testing in high- and low-income countries: Closing the diagnostic and treatment gap. Annual Meeting of American Epilepsy Society (AES). Los Angeles, USA: American Epilepsy Society (AES); 2024. https://aesnet.org/abstractslisting/international-study-of-epilepsy-genetic-testing-in-high--and-low-income-countries-closing-the-diagnostic-and-treatment-gap.
79. Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. *Cell*. 2019;177(1):26–31. doi:10.1016/j.cell.2019.02.048
80. Pasquier L, Minguet G, Moisdon-Chataigner S, et al. How do non-geneticist physicians deal with genetic tests? A qualitative analysis. *Eur J Hum Genet*. 2022;30(3):320–331. doi:10.1038/s41431-021-00884-z
81. Zhong A, Darren B, Loiseau B, et al. Ethical, social, and cultural issues related to clinical genetic testing and counseling in low- and middle-income countries: A systematic review. *Genet Med*. 2021;23(12):2270–2280. doi:10.1038/s41436-018-0090-9
82. Adamu A, Chen R, Li A, Xue G. Epilepsy in Asian countries. *Acta Epileptol*. 2023;5(1):25. doi:10.1186/s42494-023-00136-1
83. Samia P, Sahu JK, Ali A, et al. Telemedicine for Individuals with epilepsy: Recommendations from the International League Against Epilepsy Telemedicine Task Force. *Seizure*. 2023;106:85–91. doi:10.1016/j.seizure.2023.02.005
84. D'Gama AM, Mulhern S, Sheidley BR, et al. Evaluation of the feasibility, diagnostic yield, and clinical utility of rapid genome sequencing in infantile epilepsy (Gene-STEPS): An international, multicentre, pilot cohort study. *Lancet Neurol*. 2023;22(9):812–825. doi:10.1016/S1474-4422(23)00246-6
85. Ademuwagun IA, Adam Y, Rotimi SO, et al. Exome sequencing in Nigerian children with early-onset epilepsy syndromes. *Epilepsia Open*. 2025;10(1):222–232. doi:10.1002/epi4.13106

Advances in immune checkpoint inhibitor therapy for microsatellite instability-high (MSI-H) and mismatch repair-deficient (dMMR) colorectal cancer: Insights from the tumor immune microenvironment

Yu Zhao[A,C,D], Li Li[B,C,F], Bo Zhang[B,C,E], Bin Li[B,E,F], Wenjing Zhou[B,C,F]

Qingdao Hiser Hospital Affiliated of Qingdao University (Qingdao Traditional Chinese Medicine 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(5):929–938

Address for correspondence
Bin Li
E-mail: 19947683268@163.com

Funding sources
None declared

Conflict of interest
None declared

Received on April 3, 2025
Reviewed on June 9, 2025
Accepted on July 11, 2025

Published online on May 27, 2026

Abstract

This review systematically analyzes the relationship between the immune microenvironment characteristics of microsatellite instability-high (MSI-H) or deficient mismatch repair (dMMR) colorectal cancer (CRC) and the efficacy of immune checkpoint inhibitors (ICIs). The article emphasizes that this tumor subtype has a high mutation burden, abundant neoantigens, and significant immune cell infiltration, explaining its high sensitivity to immunotherapy, while also pointing out that some patients exhibit primary non-response or subsequent resistance. Based on single-cell and spatial omics, as well as multi-omics integration analyses, the authors reveal the complexity and heterogeneity of key immune cell subpopulations, spatial distribution, and resistance mechanisms (such as abnormal Janus kinase/signal transducer and activator of transcription (JAK/STAT) pathways, human leukocyte antigen (HLA) loss, and metabolic reprogramming, and propose the necessity of multi-time-point dynamic monitoring and multimodal combination therapy. The study underscores that, in the future, standardized data integration and the establishment of artificial intelligence (AI) prediction models will be required to facilitate the implementation of precise, individualized immunotherapy strategies, thereby further improving clinical efficacy.

Key words: immune checkpoint inhibitors, tumor immune microenvironment, MSI-H/dMMR, single-cell and spatial transcriptomics

Cite as
Zhao Y, Li L, Zhang B, Li B, Zhou W. Advances in immune checkpoint inhibitor therapy for microsatellite instability-high (MSI-H) and mismatch repair-deficient (dMMR) colorectal cancer: Insights from the tumor immune microenvironment. *Adv Clin Exp Med*. 2026;35(5):929–938. doi:10.17219/acem/208178

DOI
10.17219/acem/208178

Highlights

- Microsatellite instability-high (MSI-H)/deficient mismatch repair (dMMR) colorectal cancer (CRC) exhibits high tumor mutational burden (TMB), neoantigen load, and immune infiltration, explaining its strong response to immune checkpoint inhibitors (ICIs).
- A subset of MSI-H/dMMR CRC patients develops primary or acquired immunotherapy resistance driven by immune escape mechanisms.
- Single-cell and multi-omics analyses reveal immune heterogeneity, spatial complexity, and resistance pathways in CRC.
- Resistance mechanisms include Janus kinase/signal transducer and activator of transcription (JAK/STAT) pathway alterations, human leukocyte antigen (HLA) loss, and tumor metabolic reprogramming.
- Precision immunotherapy strategies require dynamic monitoring, combination regimens, AI-based prediction models, and standardized multi-omics integration.

Introduction

Colorectal cancer (CRC) is one of the most common malignant tumors worldwide and a leading cause of cancer-related death. In 2020, it was estimated that there were approx. 1.9 million new CRC cases and 930,000 deaths globally.[1] With population aging and the increasing prevalence of unhealthy lifestyles, the incidence of CRC in low- and middle-income countries has been rising rapidly. Although screening and treatment have made significant progress, the prognosis for advanced CRC remains poor; the 5-year survival rate for stage IV disease is only about 14%. Particularly in metastatic CRC, conventional treatments (surgery, chemotherapy, and biologically targeted therapies) often fail to achieve durable responses, highlighting the urgent need for new therapeutic strategies.[2]

Microsatellite instability-high (MSI-H) or deficient mismatch repair (dMMR) CRC accounts for about 15% of all CRC cases. Due to defects in the DNA mismatch repair pathway, these tumors accumulate a large number of mutations, thereby generating abundant neoantigens and resulting in a high tumor mutation burden (TMB).[3] MSI-H/dMMR CRC is usually characterized by rich immune cell infiltration (a "hot" tumor) and displays strong immunogenicity. In contrast, microsatellite-stable (MSS) or mismatch repair-proficient (pMMR) CRC exhibits relatively low immune infiltration and is considered an immune-tolerant "cold" tumor.[4] This difference in the immune microenvironment partly explains the marked sensitivity of MSI-H/dMMR CRC to immune checkpoint inhibitors (ICIs).

Since 2017, ICIs have been approved for the treatment of MSI-H/dMMR metastatic CRC. Phase III clinical trials, such as KEYNOTE-177, have shown that PD-1 inhibitors (e.g., pembrolizumab) can significantly prolong progression-free survival (PFS) in patients with MSI-H mCRC compared with chemotherapy.[5] However, approx. 40–60% of MSI-H/dMMR CRC patients still exhibit primary non-response or subsequently develop acquired resistance to ICIs. In some cases, immune therapy-related hyperprogression (HPD), characterized by unexpectedly accelerated tumor growth, has been observed.[6] Therefore, an in-depth analysis of the unique immune microenvironment characteristics of MSI-H/dMMR CRC and their relationship to therapeutic efficacy is of great significance for improving immunotherapy outcomes.

Objectives

This review aims to systematically summarize research progress on the tumor immune microenvironment (TME) of MSI-H/dMMR CRC since 2018, focusing on: 1) the roles of various TME components (such as TMB, neoantigens, immune cells, fibroblasts, and stroma); 2) the heterogeneity of key immune cell subpopulations and their responses to treatment; 3) potential mechanisms of immunotherapy resistance and HPD; 4) the application of new technologies, such as single-cell sequencing and spatial transcriptomics; 5) novel biomarkers identified through multi-omics integration; and 6) current challenges and future trends. By reviewing these cutting-edge advances, we aim to provide a reference for developing individualized immunotherapy strategies and to emphasize the importance of large-scale multi-omics research and multidisciplinary collaboration.

Materials and methods

Databases including PubMed, Web of Science, Embase, and Scopus were searched up to March 2025, primarily using the following terms and their combinations: "MSI-H," "dMMR," "colorectal cancer," "immune microenvironment," "immune checkpoint inhibitors," "immunotherapy," "tumor mutational burden," "single-cell sequencing," and "spatial transcriptomics," among others. To ensure comprehensiveness, we also performed a "snowball" tracing search of the reference lists of the included studies. After

exporting the literature, duplicate records were first removed using EndNote 20 (Clarivate, London, UK). Subsequently, 2 researchers (Y.Z. and L.L.) independently conducted a preliminary screening of titles and abstracts. Any discrepancies were resolved through discussion with a 3rd researcher (B.Z.). inclusion criteria were as follows: 1) studies relevant to immune microenvironment characteristics or immunotherapy in MSI-H/dMMR colorectal cancer; 2) study types including original research (basic or clinical), retrospective or prospective cohort studies, or high-quality reviews; and 3) complete experimental or clinical data supporting the research objectives. The exclusion criteria were as follows: 1) irrelevance to the topic of this review; 2) non-peer-reviewed literature or conference abstracts with insufficient information; and 3) duplicate publications or redundant data. Uniform data extraction was performed for the final included literature, encompassing: 1) study population and sample size; 2) MSI-H/dMMR classification and detection methods; 3) methods for immune microenvironment analysis (such as immunohistochemistry (IHC), single-cell RNA sequencing (scRNA-seq), spatial transcriptomics, etc.); and 4) main findings and clinical significance. Considering the level of evidence of the included literature, we comprehensively evaluated the reliability and applicability of the findings. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart outlines the process from the initial retrieval to the final inclusion of the literature (Fig. 1).

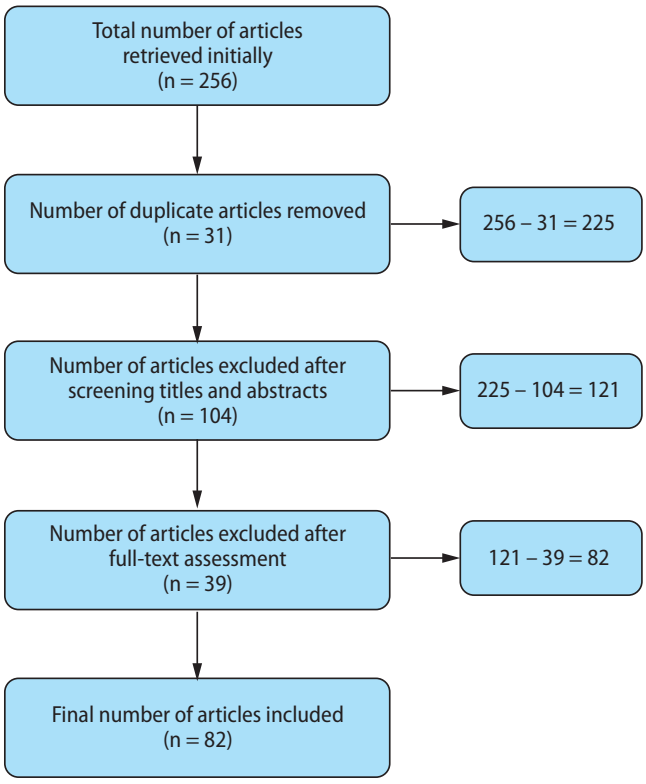

Fig. 1. Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) flowchart

The core role of the tumor immune microenvironment in MSI-H/dMMR colorectal cancer

High tumor mutation burden and neoantigen generation

Due to defects in DNA mismatch repair, MSI-H/dMMR CRC produces frameshift and insertion/deletion mutations scattered throughout the genome, generating a large number of neoantigens and resulting in an extremely high TMB, often several times higher than that of MSS tumors.[7] A high TMB increases the opportunities for antigen presentation by MHC class I molecules, thereby stimulating the clonal expansion of effector T cells. Studies have shown a positive correlation between the abundance of neoantigens and the level of tumor-infiltrating lymphocytes (TILs) in MSI-H/dMMR tumors, and this "antigen flood" phenomenon can activate the immune system. In addition, a high TMB is often accompanied by the upregulation of MHC class II molecules, which is conducive to $CD4^+$ T cell-mediated antitumor immunity.[8] Therefore, TMB and neoantigen burden are regarded as key markers for predicting the efficacy of ICIs in MSI-H/dMMR CRC.[9] However, TMB is not absolutely reliable; a small number of MSS CRC cases can also exhibit TMB levels similar to those of MSI-H tumors,[10] indicating the need for comprehensive evaluation in conjunction with other microenvironmental factors.

Microenvironment constituent elements and synergistic effects

The TME of MSI-H/dMMR CRC is composed of multiple cell types and factors that influence the intensity of the immune response and therapeutic outcomes. Immune cells are the core component: MSI-H/dMMR tumors are enriched with large numbers of $CD8^+$ cytotoxic T cells and Th1-type $CD4^+$ T cells, accompanied by high expression of cytokines such as interferon gamma (IFN-γ), thereby presenting an "inflammatory" state.[11] The proportion of regulatory T cells (Tregs) is relatively decreased, promoting antitumor immunity.[12] Natural killer (NK) cells and γδ T cells also play important roles; MSI-H CRC lacking human leukocyte antigen class I (HLA-I) can mediate ICI sensitivity through γδ T cells.[13]

Tumor-associated macrophages (TAMs) and myeloid-derived suppressor cells (MDSCs) have a double-edged effect: M1 macrophages secrete interleukin (IL)-12 and other factors to enhance anticancer immunity,[14] whereas M2 macrophages enriched in some MSI-H CRCs express PD-L1, thereby inhibiting T cell function. Cancer-associated fibroblasts (CAFs) and endothelial cells constitute the tumor stroma and vascular network; CAFs can thicken the stroma and obstruct T cell infiltration via pathways

such as transforming growth factor beta (TGF-β), while high endothelial microvasculature facilitates lymphocyte entry into the tumor bed.[15] Immunosuppressive molecules (such as IL-10 and vascular endothelial growth factor (VEGF)) and metabolic factors (such as lactate) also continuously shape the immune landscape.

Clinical significance

The immune microenvironment features of MSI-H/dMMR CRC determine its high sensitivity to ICI treatment: A large number of $CD8^+$ TILs and high PD-1/PD-L1 expression constitute the basis of immunotherapy.[16] In a study involving more than 450 CRC cases, the TIL density in the MSI-H subgroup was significantly higher than that in the MSS subgroup, and high TIL density was associated with longer survival.[17] For MSI-H/dMMR CRC patients not receiving immunotherapy, abundant immune infiltration is associated with a slightly better prognosis than in MSS CRC, with a lower recurrence risk in stages I–II.[18]

Multiple studies have also indicated better clinical outcomes in MSI-H/dMMR CRC patients in stages I–III, including significantly reduced recurrence risk and prolonged disease-free survival (DFS) and overall survival (OS),[19] which are closely associated with high levels of $CD8^+$ T cell infiltration, Th1-type immune responses, and strong neoantigen-induced immune effects within the TME. At the metastatic stage, only about 5% of mCRC cases remain MSI-H, and this advantage disappears; however, the enriched immune microenvironment still allows some patients to significantly benefit from immune checkpoint inhibitor treatment.[20]

Timely identification of MSI-H/dMMR status is crucial for clinical decision-making, and MSI testing has become routine in the molecular diagnosis of CRC, not only for Lynch syndrome screening but also as a companion diagnostic for immunotherapy.[21] These studies further support the close association between the unique immune microenvironment of MSI-H/dMMR CRC and patient prognosis, providing a reference for the clinical identification of potential beneficiaries.

Key immune cell subpopulations and the heterogeneity of treatment response

$CD8^+$ cytotoxic T cells are the main antitumor force in MSI-H/dMMR CRC, and the intratumoral density of $CD8^+$ T cells is positively correlated with ICI efficacy and survival.[22] PD-1 inhibitors can restore the killing function of exhausted $CD8^+$ T cells. Postoperative analyses indicate that *CD8A* gene (a T cell cytotoxic molecule) expression is significantly higher in ICI responders than in non-responders. Th1-type $CD4^+$ T helper cells promote the proliferation and activation of $CD8^+$ T cells by secreting IL-2 and IFN-γ, which are also crucial in the MSI-H microenvironment.[23]

In contrast, Tregs ($FoxP3^+$) infiltrate at relatively lower levels in MSI-H CRC, which benefits immune efficacy.[24] It should be noted that there is heterogeneity in T cell subpopulation composition among different patients; some MSI-H tumors are "indifferent," exhibiting few $CD8^+$ T cells and a higher proportion of Tregs, resulting in no initial response to ICIs.[25]

Atypical T cells: The role of γδ T cells

The γδ T cells combine both innate and adaptive immune features and possess MHC-nonrestricted cytotoxic activity. In CRC, γδ T cells are mainly of the $V\delta1^+$ phenotype, and their degree of infiltration is often associated with an improved prognosis.[26] In MSI-H CRC with HLA-I deficiency, γδ T cells may replace $CD8^+$ T cells in performing immune surveillance.

Yu et al. reported that the response of HLA-I-negative dMMR CRC to ICIs depends on the function of tumor-infiltrating γδ T cells, suggesting that they serve as a "replacement" cell type in the immune response of MSI-H CRC, especially in immune escape phenotypes.[27]

Myeloid cells: The heterogeneity driven by MDSCs and macrophages

MDSCs and TAMs often exert immunosuppressive effects. The number of MDSCs in the peripheral blood and tumor tissue of MSI-H CRC patients is positively correlated with resistance to ICI treatment.[28] Tumor-associated macrophages have 2 polarized states: M1 and M2. M1 macrophages are pro-inflammatory and antitumor, whereas M2 macrophages secrete IL-10, TGF-β, and other factors that inhibit immune responses.[29] Compared with MSS CRC, MSI-H CRC has more M1 macrophages and fewer M2 macrophages at the invasive margin,[30] which may explain its greater sensitivity to immunotherapy. However, in MSI-H patients with ICI resistance, TAMs can repolarize toward the M2 phenotype and upregulate immunosuppressive molecules such as PD-L1 and indoleamine 2,3-dioxygenase (IDO).[31] Therefore, combined strategies targeting MDSCs and M2 TAMs (e.g., CSF-1R or IDO inhibitors plus PD-1 antibodies) are being explored to improve efficacy in refractory MSI-H CRC patients.[32]

Spatial distribution characteristics: Differences in immune cell location within the tumor

In addition to quantity, the spatial distribution of immune cells in the TME also determines their function. In MSI-H/dMMR CRC, TILs often penetrate deep into the tumor center to form an "immune-hot" zone, whereas in MSS CRC,

TILs are mostly confined to the invasive margin, presenting an "immune exclusion" state.[33] High-resolution spatial omics comparing the tumor–stroma boundary in dMMR and pMMR CRC reveal that the boundary in dMMR CRC is enriched with LAMP3$^+$ dendritic cells and CXCL13$^+$ T cell clusters, facilitating the formation of effective immunological synapses, whereas in pMMR CRC, CAFs and collagen deposition at the boundary hinder T cell infiltration.[34] Pelka et al. reported that MSI-H tumors contain a multicellular immune hub composed of densely packed antigen-presenting cells and effector T cells adjacent to surrounding tumor cells, which is conducive to immune attack, whereas MSS tumors often lack this structure.[35]

Mechanisms of resistance and hyperprogression

Regulation of immune checkpoints and signaling pathways

Although MSI-H/dMMR CRC is generally sensitive to ICIs, some patients exhibit primary or acquired resistance, with various underlying mechanisms. Abnormalities in the PD-1/PD-L1 axis are common factors; Chen et al. found that in MSI-H CRC, PD-L1 is primarily expressed by macrophages rather than by tumor cells, which may reduce the efficacy of ICIs.[36] The classic JAK/STAT pathway is crucial for antitumor immunity, and loss-of-function mutations in *JAK1* or *JAK2* render tumors unresponsive to IFN-γ, allowing them to escape immune surveillance. Although their incidence is only about 4–7% in MSI-H CRC,[37] once they occur, they are often directly related to ICI resistance.[38]

Hyperactivation of the Wnt/β-catenin pathway can also drive immune evasion; excessive Wnt signaling suppresses chemokine production and hinders the recruitment of dendritic cells and T cells.[39] In MSS CRC, Wnt signaling promotes CAFs to secrete more TGF-β and exclude T cells. If MSI-H CRC also develops secondary Wnt abnormalities, it may transform into an immune-indifferent phenotype, leading to resistance.[40] Therefore, combination therapies targeting key signaling nodes such as JAK or Wnt may overcome some forms of resistance.

Definition of hyperprogression and exploration of mechanisms

Hyperprogression refers to the rare but abnormally accelerated disease progression following ICI treatment. There is no unified standard, but it is generally defined by a significant increase in tumor growth rate on imaging or rapid clinical deterioration.[41] Although HPD is rare in MSI-H/dMMR tumors, it is not nonexistent. Zhou et al. reported a case of dMMR small intestinal cancer exhibiting HPD after treatment with a PD-1 antibody, suggesting that high lactate dehydrogenase levels, extremely low TMB and TIL infiltration, and abnormally activated TGF-β signaling might be related.[42] In other cancers, MDM2/MDM4 amplification and epidermal growth factor receptor (EGFR) mutations have been associated with an increased risk of HPD, presumably by promoting cell-cycle progression and inhibiting apoptosis, thereby accelerating tumor growth in an immunosuppressive environment. The mechanism of HPD in MSI-H CRC is not yet clear; however, given that some CRC cases harbor MDM2 amplification and EGFR mutations,[43] clinicians should remain vigilant. For cases with rapid progression, it may be prudent to test for these genes and adjust treatment accordingly.

Antigen presentation defects and HLA loss

Although MSI-H CRC is rich in neoantigens, if the antigen presentation pathway is compromised, the immune system may still "turn a blind eye." β_2-microglobulin (B2M) is an essential component of MHC-I molecules, and its gene mutations account for approx. 24% of MSI-H CRC cases, resulting in almost no HLA-I antigen presentation in tumor cells, thus evading cytotoxic T lymphocyte (CTL) recognition. In addition, mutations in HLA-A/B/C and their transcriptional activator NLRC5 are also common. A study by Kloor et al. showed that 72% of MSI-H CRC harbor at least 1 genetic alteration in the HLA-I antigen presentation pathway, with more than half being loss-of-function mutations.[44]

This indicates that most MSI-H CRC tumors undergo immune selection pressure during evolution, tending to "shut down" antigen presentation to escape immune surveillance. For such tumors, PD-1 monotherapy is limited, necessitating combination therapy with CTLA-4 inhibitors or other immune strategies to activate NK or γδ T cells and other nonconventional immune pathways. Current studies such as COSMIC-312 are evaluating dual checkpoint blockade (PD-1 + CTLA-4) in MSI-H CRC, which is expected to partially overcome resistance caused by antigen presentation defects.[45]

Metabolic reprogramming-mediated immunosuppression

Tumor metabolism and the immune microenvironment influence each other, and metabolic reprogramming can promote immune evasion. Tryptophan (Trp) metabolism is a typical example: tumor cells and MDSCs overexpress IDO1, breaking down large amounts of Trp into kynurenine (Kyn), leading to Trp depletion and Kyn accumulation, which in turn block the T-cell cycle and induce Treg expansion.[46] Research has shown that high IDO1 expression is associated with primary resistance of MSI-H CRC to PD-1 inhibitors.[47] Clinical phase I/II trials have attempted to combine the IDO inhibitor BMS-986205 with PD-1 antibodies for the treatment of MSS CRC, and although

they did not meet the primary endpoint, subgroup analyses showed some efficacy,[48] indicating the value of further exploration of metabolic targets. Lactate is also an important metabolic factor: MSI-H/dMMR CRC undergoes high glycolysis, producing large amounts of lactate, which reduces the local pH, inhibits T- and NK-cell activity, and recruits M2-type macrophages. Sun et al. noted that lactate induces dendritic cell tolerance, reduces IL-12 secretion, and weakens CTL function. Higher lactate dehydrogenase A (LDHA) expression in pMMR CRC than in dMMR CRC is considered a marker of an immunosuppressive environment.[49] Therefore, inhibitors targeting IDO1 or LDHA may improve the TME and enhance the efficacy of ICIs. Preclinical models have shown that inhibiting LDHA can restore T cell activity and enhance the effect of PD-1 antibodies.[50]

Advances in single-cell sequencing and spatial transcriptomics

Single-cell RNA sequencing for heterogeneity analysis

Traditional bulk tissue sequencing often masks differences among microenvironmental cells, whereas scRNA-seq can delineate gene expression profiles of different cell types at the single-cell level. In recent years, multiple studies have applied scRNA-seq to CRC, revealing subtle differences in immune microenvironment composition between MSI-H and MSS tumors. Wu et al. performed scRNA-seq on 6 cases of MSI-H mCRC (3 resistant to ICIs and 3 sensitive), analyzing over 56,000 cells.[51] The results showed that the greatest differences between the responsive and resistant groups lay in the proportions of $CD8^+$ T cells and monocytes: Patients in the responsive group had higher $CD8^+$ T cell infiltration, whereas resistant patients exhibited increased proportions of immunosuppressive monocytes, implying that myeloid immunity interferes with ICI efficacy. The scRNA-seq can also uncover novel cell subpopulations that are not distinguishable using traditional markers. They have identified a cluster of $CD8^+$ T cells highly expressing LAG3 and PDCD1 that were enriched in the resistant group, indicating the potential for combination therapy with LAG-3 inhibitors. Another study incorporating T cell receptor sequencing found that some MSS CRCs still possess numerous memory $CD8^+$ T cell clones that, if activated, could potentially yield a response to immunotherapy.[52]

Spatial transcriptomics for visualizing spatial structure

Although scRNA-seq reveals cellular diversity, it loses spatial information. Spatial transcriptomics (10× Visium, Slide-seq, Stereo-seq, etc.) can measure location-specific gene expression in tissue sections. Feng et al. used Stereo-seq to analyze 5 untreated pMMR and 5 post-PD-1-treated dMMR CRC samples,[53] constructing a digital tumor atlas and identifying an "immune-activated" spatial subtype enriched with $LAMP3^+$ dendritic cells and $CXCL13^+$ T cells at the tumor–stroma interface in dMMR CRC, which was positively correlated with ICI response. In contrast, pMMR tumors presented more tumor cell–fibroblast hybrid clusters that isolated immune cells. Another study employed Visium to compare mCRC tissues before and after neoadjuvant immunotherapy,[54] showing that PD-1 antibodies facilitated deeper T-cell infiltration at the tumor boundary in dMMR CRC and reduced CAF activity, transforming it into an "immune-hot" central state.[55] The "immune hub" reported by Pelka et al.[35] was also identified through spatial analysis of immune aggregates within the tumor.

Advantages and challenges of new techniques

Single-cell and spatial omics provide unprecedented high resolution, detecting rare cell subpopulations without preset markers and enabling dynamic monitoring of multiple time point samples to track immune microenvironment remodeling. Challenges include high data complexity, requiring robust bioinformatics and rigorous statistical methods to avoid false positives or over-interpretation; elevated costs and a need for fresh tissue, posing a higher threshold for most hospitals; and limited sample sizes that may lack representativeness, leaving the question of broader applicability to further validation.

Multi-omics integration and novel biomarkers

Colorectal cancer involves multi-level changes in the genome, transcriptome, proteome, metabolome, and microbiome. A single-dimensional biomarker is often insufficient to comprehensively predict treatment outcomes, so multi-omics integration has garnered increasing attention.[56] For MSI-H/dMMR CRC, combining genomic mutations (TMB, HLA mutations), transcriptomic data (immune gene expression), proteomic data (immune infiltration scores), and microbiome data can provide a more complete picture of the tumor immune landscape.[57] Cheng et al. conducted a multi-omics study on MSI-H/dMMR advanced gastrointestinal tumors, collecting 16S microbiome sequencing data, blood metabolomics, and cytokine profiles, together with tumor molecular analyses. They found that responders were enriched in short-chain fatty acid-producing bacteria (e.g., *Akkermansia muciniphila*) and had higher levels of Trp and other metabolites in the blood, suggesting that integrating microbiome–metabolism–immune signals can form a comprehensive predictive model.[58] An immune gene expression profile (GEP) score, which weights and sums the expression of multiple immune-related genes,

can effectively predict PD-1 antibody efficacy[59]; in MSI-H CRC patients, a high GEP score is often associated with a durable response. The Immunoscore, which quantitatively assesses $CD3^+$ and $CD8^+$ T cell density in both the tumor center and the invasive margin via immunohistochemistry, outperforms TNM staging in predicting the prognosis of localized CRC.[60] In the context of immunotherapy, MSI-H mCRC with a high Immunoscore is more likely to benefit from ICIs. The serum kynurenine/tryptophan ratio (Kyn/Trp) reflects IDO activity; Cheng et al. found that a high Kyn/Trp ratio was significantly associated with early progression on immunotherapy and may be used to identify patients with primary resistance. The composition of the gut microbiome affects ICI efficacy[61]; for instance, increased abundance of *A. muciniphila* and *Bifidobacterium* is linked to ICI response in multiple tumor types.[62] In CRC, some researchers have constructed a "probiotic index" to assess the abundance of beneficial bacteria, finding that MSI-H CRC patients with a higher index respond better to PD-1 antibodies. The NICHE-2 trial applied neoadjuvant nivolumab plus ipilimumab in locally advanced (stage III) dMMR CRC and collected tumor samples before and after treatment for multi-omics analysis, showing that most patients achieved a pathological complete response. Subsequent transcriptome and T-cell receptor omics analyses will further elucidate molecular features related to efficacy.[63] The DURVA-TREM trial is evaluating the efficacy of durvalumab (anti–PD-L1) combined with tremelimumab (anti-CTLA-4) in MSI-H mCRC, using ctDNA and peripheral immune profiling to identify predictive factors.[64]

Gut microbiome interventions are also under exploration, where probiotics or fecal microbiota transplantation combinfed with ICIs are being used in MSS CRC to potentially convert it into a "hot" tumor.[65] In a study of 20 MSS CRC patients, anti-PD-1 therapy plus fecal microbiota transplantation was accompanied by pre- and post-treatment transcriptomic and microbiome analysis. Some previously "cold" tumors showed upregulated IFN-γ signaling and increased TILs.[66] Proteomics can capture post-translational modifications and protein interaction networks, thereby identifying key regulatory proteins associated with immune escape or therapeutic resistance. Metabolomics helps to elucidate the roles of lactate, Kyn, and other metabolites in immune suppression or immune activation processes. Existing studies have utilized proteomics to identify differentially expressed or modified signaling molecules, such as phosphorylated JAK/STAT proteins and critical immune checkpoint proteins, providing more precise evidence for multi-pathway combination therapies and the discovery of new therapeutic targets[67]; researchers have also used metabolomics approaches to reveal the close association between the Kyn–Trp pathway, T-cell exhaustion, and immunosuppressive TME.[68] In combination with genomic, transcriptomic, and microbiomic data, proteomics and metabolomics analyses can further refine the comprehensive landscape of the TME, facilitating the identification of biomarkers with greater clinical utility and guiding personalized immunotherapy strategies.

Unresolved key issues and future perspectives

Dynamic monitoring and multi-time point sampling

The TME continually evolves during disease progression and treatment. Most studies based on single-time point analyses have difficulty capturing its dynamic changes. Larger-scale prospective studies need to collect tissues from the same patient at multiple time points before, during, and after treatment through longitudinal cohorts, combined with dynamic indicators such as peripheral T cell subsets and ctDNA to monitor shifts in the immune landscape throughout the course of ICIs.[69] Neoadjuvant immunotherapy can perform multi-omics analyses before, during, and after treatment to track changes in immune infiltration from baseline to pathological remission; minimally invasive liquid biopsies can capture the real-time status of peripheral blood immune cells.[70] Although multi-time point sampling has ethical and feasibility concerns, patient-derived xenograft (PDX) mouse models with humanized immune systems can only partially simulate tumor evolution.[71]

Exploration of combination therapy strategies

Although single checkpoint blockade in MSI-H/dMMR CRC shows significant effectiveness, improvements are still needed for resistant and MSS populations. Multimodal combination strategies remain under continuous exploration. For instance, PD-1 + CTLA-4 can increase response rates in metastatic MSI-H CRC,[72] and other studies incorporate LAG-3 or TIGIT antibodies to further activate exhausted T cells. Radio-chemotherapy (e.g., oxaliplatin) can induce immunogenic cell death and enhance immune stimulation in the TME[73]; small-molecule targeting (JAK/STAT agonists or Wnt inhibitors) can also alleviate immunosuppression.[74] Personalized neoantigen vaccines have shown promise in melanoma,[75] though their clinical application in MSI-H CRC still lacks sufficient evidence; TIL or TCR-T cell therapies for MSI-H CRC are under clinical investigation (NCT03412877), and their actual clinical benefits require further evidence support.

Data standardization and multidisciplinary integration

Multi-omics and high-throughput technologies generate massive multidimensional data, urgently requiring

unified data processing standards and multidisciplinary collaboration.[76] Employing uniform formats and metadata descriptions (e.g., clinical information, sequencing parameters) facilitates subsequent meta-analyses. Artificial intelligence (AI) can extract patterns from multi-level data to predict individual treatment responses,[77] but close cooperation with tumor biologists and clinical experts is needed to ensure practicality. Some studies have attempted to integrate radiomics, pathological images, and gene expression with deep learning models to predict MSI status,[78] with feasibility and accuracy still needing further validation.

Controversial issues and future research directions

The value of anti-EGFR therapy in the MSI-H setting remains controversial. Conventional wisdom holds that RAS/BRAF wild-type CRC is limited to left-sided MSS cases for use, but recent reports show significant tumor shrinkage with cetuximab in patients who are MSI-H and RAS/BRAF wild-type,[79,80] though the applicable population still requires more sufficient multicenter validation. Other studies suggest that MSI-H/dMMR tumors under high mutation and strong immune pressure may more easily develop immune evasion, with some patients experiencing relapse after immunotherapy and loss of dMMR features,[81] potentially related to immune clonal selection or inadequate treatment. Additionally, the impact of the tumor microbiome (e.g., *Fusobacterium* spp.) on the TEM and mechanisms maintaining immune memory exhibit variations, and specific intervention strategies and clinical benefits currently lack large-sample evidence.[82] Further in-depth prospective studies and multi-dimensional validations may advance the refinement of personalized therapeutic strategies.

Limitations of the study

This review still has limitations in both data collection and scope. It relies primarily on published literature without conducting a meta-analysis, and may be influenced by study biases (such as differences in study populations and technical platforms). Some mechanisms (e.g., the potential mechanisms of HPD and the value of new biomarkers) are based on limited evidence and require further experimental validation. Since MSI-H/dMMR CRC accounts for a relatively small proportion of cases and single-cell and spatial omics studies have limited sample sizes, caution is needed when applying these conclusions to larger populations. Due to space constraints, an in-depth discussion of neoadjuvant immunotherapy and insights from other MSI-H solid tumors was not possible. Readers should critically integrate the latest developments with each patient's specific circumstances.

Conclusions

MSI-H/dMMR CRC is characterized by a high mutation burden and abundant immune infiltration, conferring a relatively favorable prognosis and providing ideal conditions for immune checkpoint inhibitor therapy. In recent years, PD-1 monoclonal antibodies have shown advantages in first-line treatment, and dual blockade of PD-1/CTLA-4 has further enhanced clinical benefits. However, some patients still exhibit no response or develop resistance. This paper has outlined the immune microenvironment characteristics of MSI-H/dMMR CRC, including the strong immune response driven by a high TMB and neoantigen burden, as well as the mechanisms of immune escape caused by HLA mutations, JAK pathway abnormalities, immunosuppressive cell infiltration, and metabolic reprogramming. Single-cell and spatial multi-omics technologies are helping us more accurately characterize immune cell types and distributions, clarifying individual differences in treatment outcomes.

Use of AI and AI-assisted technologies

Not applicable.

ORCID iDs

Yu Zhao https://orcid.org/0009-0001-8298-0959
Li Li https://orcid.org/0009-0001-0106-989X
Bo Zhang https://orcid.org/0009-0002-3507-8038
Wenjing Zhou https://orcid.org/0009-0004-8055-1255
Bin Li https://orcid.org/0009-0004-5411-1141

References

1. Morgan E, Arnold M, Gini A, et al. Global burden of colorectal cancer in 2020 and 2040: Incidence and mortality estimates from GLOBOCAN. *Gut*. 2023;72(2):338–344. doi:10.1136/gutjnl-2022-327736
2. American Cancer Society (ACS). Survival Rates for Colorectal Cancer. Atlanta, USA: American Cancer Society (ACS); 2023. https://www.cancer.org/cancer/types/colon-rectal-cancer/detection-diagnosis-staging/survival-rates.html. Accessed March 15, 2025.
3. Ganesh K, Stadler ZK, Cercek A, et al. Immunotherapy in colorectal cancer: Rationale, challenges and potential. *Nat Rev Gastroenterol Hepatol*. 2019;16(6):361–375. doi:10.1038/s41575-019-0126-x
4. Mlecnik B, Bindea G, Angell HK, et al. Integrative analyses of colorectal cancer show immunoscore is a stronger predictor of patient survival than microsatellite instability. *Immunity*. 2016;44(3):698–711. doi:10.1016/j.immuni.2016.02.025
5. André T, Shiu KK, Kim TW, et al. Pembrolizumab in microsatellite-instability–high advanced colorectal cancer. *N Engl J Med*. 2020;383(23):2207–2218. doi:10.1056/nejmoa2017699
6. Wang Q, Shen X, Chen G, Du J. How to overcome resistance to immune checkpoint inhibitors in colorectal cancer: From mechanisms to translation. *Int J Cancer*. 2023;153(4):709–722. doi:10.1002/ijc.34464
7. Xie YH, Chen YX, Fang JY. Comprehensive review of targeted therapy for colorectal cancer. *Sig Transduct Target Ther*. 2020;5(1):22. doi:10.1038/s41392-020-0116-z
8. Haabeth OAW, Tveita AA, Fauskanger M, et al. How do $CD4^+$ T cells detect and eliminate tumor cells that either lack or express MHC class II molecules? *Front Immunol*. 2014;5:174. doi:10.3389/fimmu.2014.00174
9. Miao D, Margolis CA, Vokes NI, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat Genet*. 2018;50(9):1271–1281. doi:10.1038/s41588-018-0200-2

10. Lee WS, Yang H, Chon HJ, Kim C. Combination of anti-angiogenic therapy and immune checkpoint blockade normalizes vascular-immune crosstalk to potentiate cancer immunity. *Exp Mol Med*. 2020; 52(9):1475–1485. doi:10.1038/s12276-020-00500-y
11. Lin A, Zhang J, Luo P. Crosstalk between the MSI status and tumor microenvironment in colorectal cancer. *Front Immunol*. 2020;11:2039. doi:10.3389/fimmu.2020.02039
12. Duhen T, Duhen R, Montler R, et al. Co-expression of CD39 and CD103 identifies tumor-reactive CD8 T cells in human solid tumors. *Nat Commun*. 2018;9(1):2724. doi:10.1038/s41467-018-05072-0
13. De Vries NL, Van De Haar J, Veninga V, et al. γδ T cells are effectors of immunotherapy in cancers with HLA class I defects. *Nature*. 2023; 613(7945):743–750. doi:10.1038/s41586-022-05593-1
14. Zheng P, He J, Fu Y, et al. Engineered bacterial biomimetic vesicles reprogram tumor-associated macrophages and remodel tumor microenvironment to promote innate and adaptive antitumor immune responses. *ACS Nano*. 2024;18(9):6863–6886. doi:10.1021/acsnano.3c06987
15. Georganaki M, Van Hooren L, Dimberg A. Vascular targeting to increase the efficiency of immune checkpoint blockade in cancer. *Front Immunol*. 2018;9:3081. doi:10.3389/fimmu.2018.03081
16. Boland CR, Goel A. Microsatellite instability in colorectal cancer. *Gastroenterology*. 2010;138(6):2073–2087.e3. doi:10.1053/j.gastro.2009.12.064
17. Rozek LS, Schmit SL, Greenson JK, et al. Tumor-infiltrating lymphocytes, Crohn's-like lymphoid reaction, and survival from colorectal cancer. *J Natl Cancer Inst*. 2016;108(8):djw027. doi:10.1093/jnci/djw027
18. Luchini C, Bibeau F, Ligtenberg MJL, et al. ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: A systematic review-based approach. *Ann Oncol*. 2019; 30(8):1232–1243. doi:10.1093/annonc/mdz116
19. REACCT Collaborative; Zaborowski AM, Abdile A, Adamina M, et al. Impact of microsatellite status in early-onset colonic cancer. *Br J Surg*. 2022;109(7):632–636. doi:10.1093/bjs/znac108
20. Cervantes B, André T, Cohen R. Deficient mismatch repair/microsatellite unstable colorectal cancer: Therapeutic advances and questions. *Ther Adv Med Oncol*. 2024;16:17588359231170473. doi:10.1177/17588359231170473
21. Snowsill T, Coelho H, Huxley N, et al. Molecular testing for Lynch syndrome in people with colorectal cancer: Systematic reviews and economic evaluation. *Health Technol Assess*. 2017;21(51):1–238. doi:10.3310/hta21510
22. Wang F, Zhao Q, Wang YN, et al. Evaluation of *POLE* and *POLD1* mutations as biomarkers for immunotherapy outcomes across multiple cancer types. *JAMA Oncol*. 2019;5(10):1504. doi:10.1001/jamaoncol.2019.2963
23. Griffith OL, Spies NC, Anurag M, et al. The prognostic effects of somatic mutations in ER-positive breast cancer. *Nat Commun*. 2018; 9(1):3476. doi:10.1038/s41467-018-05914-x
24. Le Gouvello S, Bastuji-Garin S, Aloulou N, et al. High prevalence of Foxp3 and IL17 in MMR-proficient colorectal carcinomas. *Gut*. 2008;57(6):772–779. doi:10.1136/gut.2007.123794
25. Le DT, Durham JN, Smith KN, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science*. 2017;357(6349): 409–413. doi:10.1126/science.aan6733
26. Xu H, Cao C, Ren Y, et al. Antitumor effects of fecal microbiota transplantation: Implications for microbiome modulation in cancer treatment. *Front Immunol*. 2022;13:949490. doi:10.3389/fimmu.2022.949490
27. Yu X, Wang L, Niu Z, Zhu L. Controversial role of γδ T cells in colorectal cancer. *Am J Cancer Res*. 2024;14(4):1482–1500. doi:10.62347/HWMB1163
28. Yin K, Xia X, Rui K, Wang T, Wang S. Myeloid-derived suppressor cells: A new and pivotal player in colorectal cancer progression. *Front Oncol*. 2020;10:610104. doi:10.3389/fonc.2020.610104
29. Bose D, Banerjee S, Chatterjee N, Das S, Saha M, Saha KD. Inhibition of TGF-β induced lipid droplets switches M2 macrophages to M1 phenotype. *Toxicol In Vitro*. 2019;58:207–214. doi:10.1016/j.tiv.2019.03.037
30. Narayanan S, Kawaguchi T, Peng X, et al. Tumor infiltrating lymphocytes and macrophages improve survival in microsatellite unstable colorectal cancer. *Sci Rep*. 2019;9(1):13455. doi:10.1038/s41598-019-49878-4
31. Heregger R, Huemer F, Steiner M, Gonzalez-Martinez A, Greil R, Weiss L. Unraveling resistance to immunotherapy in MSI-high colorectal cancer. *Cancers (Basel)*. 2023;15(20):5090. doi:10.3390/cancers15205090
32. Pastor F, Berraondo P, Etxeberria I, et al. An RNA toolbox for cancer immunotherapy. *Nat Rev Drug Discov*. 2018;17(10):751–767. doi:10.1038/nrd.2018.132
33. Nearchou IP, Gwyther BM, Georgiakakis ECT, et al. Spatial immune profiling of the colorectal tumor microenvironment predicts good outcome in stage II patients. *NPJ Digit Med*. 2020;3(1):71. doi:10.1038/s41746-020-0275-x
34. Feng Y, Ma W, Zang Y, et al. Spatially organized tumor-stroma boundary determines the efficacy of immunotherapy in colorectal cancer patients. *Nat Commun*. 2024;15(1):10259. doi:10.1038/s41467-024-54710-3
35. Pelka K, Hofree M, Chen JH, et al. Spatially organized multicellular immune hubs in human colorectal cancer. *Cell*. 2021;184(18): 4734–4752.e20. doi:10.1016/j.cell.2021.08.003
36. Chen L, Jiang X, Li Y, et al. How to overcome tumor resistance to anti-PD-1/PD-L1 therapy by immunotherapy modifying the tumor microenvironment in MSS CRC. *Clin Immunol*. 2022;237:108962. doi:10.1016/j.clim.2022.108962
37. Ozcan M, Janikovits J, Von Knebel Doeberitz M, Kloor M. Complex pattern of immune evasion in MSI colorectal cancer. *Oncoimmunology*. 2018;7(7):e1445453. doi:10.1080/2162402x.2018.1445453
38. Zhao P, Li L, Jiang X, Li Q. Mismatch repair deficiency/microsatellite instability-high as a predictor for anti-PD-1/PD-L1 immunotherapy efficacy. *J Hematol Oncol*. 2019;12(1):54. doi:10.1186/s13045-019-0738-1
39. Goldsberry WN, Londoño A, Randall TD, Norian LA, Arend RC. A review of the role of Wnt in cancer immunomodulation. *Cancers (Basel)*. 2019;11(6):771. doi:10.3390/cancers11060771
40. Yang W, Li Z, Li Y, He W, Yan J. Transforming albumin into a Trojan horse of immunotherapy-resistant colorectal cancer with a high microsatellite instability. *ACS Nano*. 2024;18(29):19332–19344. doi:10.1021/acsnano.4c05893
41. Ferrara R, Mezquita L, Texier M, et al. Hyperprogressive disease in patients with advanced non-small cell lung cancer treated with PD-1/PD-L1 inhibitors or with single-agent chemotherapy. *JAMA Oncol*. 2018;4(11):1543. doi:10.1001/jamaoncol.2018.3676
42. Zhou W, Zhou Y, Yi C, et al. Case report: Immune and genomic characteristics associated with hyperprogression in a patient with metastatic deficient mismatch repair gastrointestinal cancer treated with anti-PD-1 antibody. *Front Immunol*. 2021;12:749204. doi:10.3389/fimmu.2021.749204
43. Zhang L, Zhu D, Jiang J, Min Z, Fa Z. The ubiquitin E3 ligase MDM2 induces chemoresistance in colorectal cancer by degradation of ING3. *Carcinogenesis*. 2023;44(7):562–575. doi:10.1093/carcin/bgad040
44. Kloor M, Michel S, Von Knebel Doeberitz M. Immune evasion of microsatellite unstable colorectal cancers. *Int J Cancer*. 2010;127(5):1001–1010. doi:10.1002/ijc.25283
45. Overman MJ, Lonardi S, Wong KYM, et al. Durable clinical benefit with nivolumab plus ipilimumab in DNA mismatch repair–deficient/microsatellite instability–high metastatic colorectal cancer. *J Clin Oncol*. 2018;36(8):773–779. doi:10.1200/jco.2017.76.9901
46. Platten M, Wick W, Van Den Eynde BJ. Tryptophan catabolism in cancer: Beyond IDO and tryptophan depletion. *Cancer Res*. 2012;72(21): 5435–5440. doi:10.1158/0008-5472.can-12-0569
47. Botticelli A, Mezi S, Pomati G, et al. Tryptophan catabolism as immune mechanism of primary resistance to anti-PD-1. *Front Immunol*. 2020; 11:1243. doi:10.3389/fimmu.2020.01243
48. Eng C, Kim TW, Bendell J, et al. Atezolizumab with or without cobimetinib versus regorafenib in previously treated metastatic colorectal cancer (IMblaze370): A multicentre, open-label, phase 3, randomised, controlled trial. *Lancet Oncol*. 2019;20(6):849–861. doi:10.1016/s1470-2045(19)30027-0
49. Sun Q, Wu J, Zhu G, et al. Lactate-related metabolic reprogramming and immune regulation in colorectal cancer. *Front Endocrinol (Lausanne)*. 2023;13:1089918. doi:10.3389/fendo.2022.1089918
50. Zhang YX, Zhao YY, Shen J, et al. Nanoenabled modulation of acidic tumor microenvironment reverses anergy of infiltrating T cells and potentiates anti-PD-1 therapy. *Nano Lett*. 2019;19(5):2774-2783. doi:10.1021/acs.nanolett.8b04296

51. Wu T, Zhang X, Liu X, et al. Single-cell sequencing reveals the immune microenvironment landscape related to anti-PD-1 resistance in metastatic colorectal cancer with high microsatellite instability. *BMC Med*. 2023;21(1):161. doi:10.1186/s12916-023-02866-y
52. Sade-Feldman M, Yizhak K, Bjorgaard SL, et al. Defining T cell states associated with response to checkpoint immunotherapy in melanoma. *Cell*. 2018;175(4):998–1013.e20. doi:10.1016/j.cell.2018.10.038
53. Feng Y, Ma W, Zang Y, et al. Spatially organized tumor-stroma boundary determines the efficacy of immunotherapy in colorectal cancer patients. *Nat Commun*. 2024;15(1):10259. doi:10.1038/s41467-024-54710-3
54. Chalabi M, Fanchi LF, Dijkstra KK, et al. Neoadjuvant immunotherapy leads to pathological responses in MMR-proficient and MMR-deficient early-stage colon cancers. *Nat Med*. 2020;26(4):566–576. doi:10.1038/s41591-020-0805-8
55. Chalabi M, Verschoor YL, Tan PB, et al. Neoadjuvant immunotherapy in locally advanced mismatch repair–deficient colon cancer. *N Engl J Med*. 2024;390(21):1949–1958. doi:10.1056/nejmoa2400634
56. Yaeger R, Chatila WK, Lipsyc MD, et al. Clinical sequencing defines the genomic landscape of metastatic colorectal cancer. *Cancer Cell*. 2018;33(1):125–136.e3. doi:10.1016/j.ccell.2017.12.004
57. Li Y, Wu X, Fang D, Luo Y. Informing immunotherapy with multi-omics driven machine learning. *NPJ Digit Med*. 2024;7(1):67. doi:10.1038/s41746-024-01043-6
58. Cheng S, Han Z, Dai D, et al. Multi-omics of the gut microbial ecosystem in patients with microsatellite-instability–high gastrointestinal cancer resistant to immunotherapy. *Cell Rep Med*. 2024;5(1):101355. doi:10.1016/j.xcrm.2023.101355
59. Ayers M, Lunceford J, Nebozhyn M, et al. IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *J Clin Invest*. 2017;127(8):2930–2940. doi:10.1172/jci91190
60. Galon J, Mlecnik B, Bindea G, et al. Towards the introduction of the 'Immunoscore' in the classification of malignant tumours. *J Pathol*. 2014;232(2):199–209. doi:10.1002/path.4287
61. Cheng WY, Wu CY, Yu J. The role of gut microbiota in cancer treatment: Friend or foe? *Gut*. 2020;69(10):1867–1876. doi:10.1136/gutjnl-2020-321153
62. Thomas AM, Fidelle M, Routy B, et al. Gut OncoMicrobiome Signatures (GOMS) as next-generation biomarkers for cancer immunotherapy. *Nat Rev Clin Oncol*. 2023;20(9):583–603. doi:10.1038/s41571-023-00785-8
63. Zhang X, Wu T, Cai X, et al. Neoadjuvant immunotherapy for MSI-H/dMMR locally advanced colorectal cancer: New strategies and unveiled opportunities. *Front Immunol*. 2022;13:795972. doi:10.3389/fimmu.2022.795972
64. André T, Lonardi S, Wong KYM, et al. Nivolumab plus low-dose ipilimumab in previously treated patients with microsatellite instability-high/mismatch repair-deficient metastatic colorectal cancer: 4-year follow-up from CheckMate 142. *Ann Oncol*. 2022;33(10):1052–1060. doi:10.1016/j.annonc.2022.06.008
65. Davar D, Zarour HM. Facts and hopes for gut microbiota interventions in cancer immunotherapy. *Clin Cancer Res*. 2022;28(20):4370–4384. doi:10.1158/1078-0432.ccr-21-1129
66. Bonneville R, Krook MA, Kautto EA, et al. Landscape of microsatellite instability across 39 cancer types. *JCO Precis Oncol*. 2017;1:PO.17.00073. doi:10.1200/po.17.00073
67. Rosenberger G, Li W, Turunen M, et al. Network-based elucidation of colon cancer drug resistance mechanisms by phosphoproteomic time-series analysis. *Nat Commun*. 2024;15(1):3909. doi:10.1038/s41467-024-47957-3
68. Wang Z, Yin M, Zhou R, Li M, Peng J, Wang Z. Kynurenine promotes the immune escape of colorectal cancer cells via NAT10-mediated ac4C acetylation of PD-L1. *Clinics (Sao Paulo)*. 2025;80:100658. doi:10.1016/j.clinsp.2025.100658
69. Dong Y, Chen Z, Yang F, Wei J, Huang J, Long X. Prediction of immunotherapy responsiveness in melanoma through single-cell sequencing-based characterization of the tumor immune microenvironment. *Transl Oncol*. 2024;43:101910. doi:10.1016/j.tranon.2024.101910
70. Cheng YK, Chen DW, Chen P, et al. Association of peripheral blood biomarkers with response to anti-PD-1 immunotherapy for patients with deficient mismatch repair metastatic colorectal cancer: A multicenter cohort study. *Front Immunol*. 2022;13:809971. doi:10.3389/fimmu.2022.809971
71. Mai Z, Lin Y, Lin P, Zhao X, Cui L. Modulating extracellular matrix stiffness: A strategic approach to boost cancer immunotherapy. *Cell Death Dis*. 2024;15(5):307. doi:10.1038/s41419-024-06697-4
72. Kasi PM, Budde G, Krainock M, et al. Circulating tumor DNA (ctDNA) serial analysis during progression on PD-1 blockade and later CTLA-4 rescue in patients with mismatch repair deficient metastatic colorectal cancer. *J Immunother Cancer*. 2022;10(1):e003312. doi:10.1136/jitc-2021-003312
73. Tesniere A, Schlemmer F, Boige V, et al. Immunogenic death of colon cancer cells treated with oxaliplatin. *Oncogene*. 2010;29(4):482–491. doi:10.1038/onc.2009.356
74. Tzeng HT, Chyuan IT, Lai JH. Targeting the JAK-STAT pathway in autoimmune diseases and cancers: A focus on molecular mechanisms and therapeutic potential. *Biochem Pharmacol*. 2021;193:114760. doi:10.1016/j.bcp.2021.114760
75. Kleine-Kohlbrecher D, Petersen NV, Pavlidis MA, et al. Abstract LB199: A personalized neoantigen vaccine is well tolerated and induces specific T-cell immune response in patients with resected melanoma. *Cancer Res*. 2023;83(8 Suppl):LB199. doi:10.1158/1538-7445.am2023-lb199
76. Kang M, Ko E, Mersha TB. A roadmap for multi-omics data integration using deep learning. *Brief Bioinform*. 2022;23(1):bbab454. doi:10.1093/bib/bbab454
77. Chong W, Zhu X, Ren H, et al. Integrated multi-omics characterization of KRAS mutant colorectal cancer. *Theranostics*. 2022;12(11):5138–5154. doi:10.7150/thno.73089
78. Echle A, Grabsch HI, Quirke P, et al. Clinical-grade detection of microsatellite instability in colorectal tumors by deep learning. *Gastroenterology*. 2020;159(4):1406–1416.e11. doi:10.1053/j.gastro.2020.06.021
79. Mauri G, Bonazzina E, Amatu A, et al. The evolutionary landscape of treatment for BRAFV600E mutant metastatic colorectal cancer. *Cancers (Basel)*. 2021;13(1):137. doi:10.3390/cancers13010137
80. Guedes A, Silva S, Custódio S, Capela A. Successful cetuximab rechallenge in metastatic colorectal cancer: A case report. *World J Clin Oncol*. 2024;15(9):1232–1238. doi:10.5306/wjco.v15.i9.1232
81. Grasso CS, Giannakis M, Wells DK, et al. Genetic mechanisms of immune evasion in colorectal cancer. *Cancer Discov*. 2018;8(6):730–749. doi:10.1158/2159-8290.cd-17-1327
82. Casasanta MA, Yoo CC, Udayasuryan B, et al. *Fusobacterium nucleatum* host–cell binding and invasion induces IL-8 and CXCL1 secretion that drives colorectal cancer cell migration. *Sci Signal*. 2020;13(641):eaba9157. doi:10.1126/scisignal.aba9157

Racial differences in ceramides, cardiovascular health and cardiovascular risk: A preliminary analysis

Mathias Lalika[1,A–F], Vlad C. Vasile[1,2,A–F], Matthew P. Johnson[3,C–F], Sharonne N. Hayes[1,A–F], Clarence Jones[4,A–F], Lisa A. Cooper[5,A–F], Christi A. Patten[6,A–F], LaPrincess C. Brewer[1,A–F]

[1] Department of Cardiovascular Medicine, Mayo Clinic, Rochester, USA
[2] Department of Laboratory Medicine and Pathology, Mayo Clinic, Rochester, USA
[3] Department of Quantitative Health Sciences, Mayo Clinic, Rochester, USA
[4] Hue-Man Partnership, Minneapolis, USA
[5] Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, USA
[6] Department of Psychiatry and Psychology, Mayo Clinic College of Medicine, Rochester, 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(5):939–944

Address for correspondence
LaPrincess C. Brewer
E-mail: Brewer.LaPrincess@mayo.edu

Conflict of interest
None declared

Acknowledgements
The authors gratefully acknowledge all study participants and the partnering church congregations of the FAITH! Heart Health+ Study for their commitment and support, particularly during the challenges of the COVID-19 pandemic. We thank the Mayo Clinic Chief Executive Office for its contributions to this work, as well as the FAITH! study team, the Mayo Clinic Clinical Research and Trials Unit, and the investigators of the Prevalence of Asymptomatic Left Ventricular Dysfunction Study for their diligent efforts in data collection. We extend special appreciation to the FAITH! Community Steering Committee for their continuous guidance and steadfast support throughout the study.

Received on February 27, 2025
Reviewed on June 23, 2025
Accepted on September 20, 2025

Published online on January 7, 2026

Cite as
Lalika M, Vasile VC, Johnson MP, et al. Racial differences in ceramides, cardiovascular health and cardiovascular risk: A preliminary analysis. *Adv Clin Exp Med*. 2026;35(5):939–944. doi:10.17219/acem/211134

DOI
10.17219/acem/211134

Abstract

Background. Plasma ceramides are recognized biomarkers of cardiovascular risk; however, racial and ethnic differences in their levels, as well as their association with cardiovascular health (CVH) among African-American populations, remain insufficiently studied.

Objectives. This study aimed to assess the association between ceramide scores and CVH, as well as atherosclerotic cardiovascular disease (ASCVD) risk, among African-American adults, and to compare ceramide scores between African-American and White adults.

Materials and methods. We conducted a secondary analysis of 2 U.S. studies including African-American and White adults. Collected data encompassed demographics, behavioral factors (e.g., diet) and clinical measures (e.g., plasma ceramide levels). Atherosclerotic cardiovascular disease risk was assessed using the American College of Cardiology/American Heart Association (ACC/AHA) 10-year pooled cohort equations, while CVH was evaluated using the American Heart Association (AHA) Life's Essential 8 (LE8) scoring system.

Results. Fifty-eight African-American adults (mean age: 54.6 years; 67.2% women) and 1,103 White adults (mean age: 64.5 years; 52.1% women) were included. Compared with White participants, African-Americans had significantly higher prevalence of obesity, hypertension, diabetes, and hyperlipidemia, but similar ASCVD risk (12.8% vs 12.6%; $p = 0.65$). No significant associations were observed between ceramide scores and either LE8 or ASCVD risk in African-Americans. Ceramide levels differed by race/ethnicity, with African-Americans showing lower concentrations of 18:0 (0.08 vs 0.10 µmol/L) and 24:1 (0.91 vs 1.17 µmol/L) species compared with White adults (both $p < 0.001$).

Conclusions. No association was observed between ceramide scores and CVH or ASCVD risk in African-American adults. Despite having a less favorable cardiometabolic profile, African-Americans exhibited lower ceramide levels than White adults. These findings suggest that ceramide scores may not accurately reflect cardiovascular risk in African-American populations.

Key words: African-American, biomarkers, cardiovascular disease, ceramides, cardiovascular risk assessment

Funding sources

This research was supported by multiple grants: the National Institutes of Health (NIH)/National Institute on Minority Health and Health Disparities (NIMHD) (grant No. 1R21 MD013490–01); the Clinical and Translational Science Awards (CTSA) program (grant No. UL1 TR000135) from the National Center for Advancing Translational Sciences (NCATS) to Mayo Clinic; the Mayo Clinic Center for Health Equity and Community Engagement in Research; and the Mayo Clinic Chief Executive Office. LaPrincess C. Brewer received additional support from the American Heart Association–Amos Medical Faculty Development Program (grant No. 19AMFDP35040005), the American Heart Association Second Century Implementation Science Award (grant No. 23SCISA1144689), NCATS (CTSA grant No. KL2 TR002379), the Bristol Myers Squibb Foundation, the Centers for Disease Control and Prevention (CDC, grant No. CDC-DP18-1817), and the NIH/NIMHD P50 Center for Chronic Disease Reduction and Equity Promotion Across Minnesota (C2DREAM) Program (grant No. P50MD017342) during the course of this work. Mathias Lalika received funding from the American Heart Association through the Research Supplement to Promote Diversity in Science (grant No. 23DIVSUP1067167). The funders had no role in study design, data collection and analysis, decision to publish, or manuscript preparation.

Highlights

- Plasma ceramide levels were lower in African-American adults compared to White adults, despite African-Americans showing higher prevalence of obesity, hypertension, diabetes, and hyperlipidemia.
- No significant association was observed between ceramide scores and ASCVD risk or cardiovascular health (Life's Essential 8) in African-American participants.
- Key ceramides, including 18:0 and 24:1, were significantly lower in African-Americans compared with White adults, suggesting racial/ethnic differences in lipid biomarkers.
- Ceramide scores may underestimate cardiovascular risk in African-Americans, highlighting the need for race-specific biomarkers or ceramide score cutoffs to improve CVD risk prediction.

Background

Although cardiovascular care has advanced, African-American adults continue to experience suboptimal cardiovascular disease outcomes compared to White adults.[1] Furthermore, while African-American adults typically exhibit higher levels of established biomarkers such as lipoprotein(a) (Lp(a)) than White adults, the strength of the association between Lp(a) and atherosclerotic cardiovascular disease (ASCVD) appears to be comparable across both groups.[2]

This evidence underscores the need to explore novel biomarkers that may enhance cardiovascular disease (CVD) risk assessment among African-American adults. Ceramides are phospholipids that play key roles in cellular functions, and some ceramide species have been associated with atherosclerotic plaque instability and adverse cardiovascular events.[3,4]

Research has demonstrated that elevated plasma ceramide species, including ceramide (d18:1/16:0), (d18:1/18:0), (d18:1/20:0), and (d18:1/24:1), are independent predictors of CVD risk.[5] These findings underscore the utility of ceramides in CVD risk assessment, even among individuals with normal lipid profiles or low-density lipoprotein cholesterol (LDL-C) concentrations below conventional treatment thresholds.[4,5] Importantly, ceramide-based risk scores have been developed for clinical application, providing efficient tools for CVD risk stratification and prevention.[4,6] However, racial and ethnic variations in ceramide concentrations, as well as their relationship to cardiovascular health (CVH) in African-American populations, remain insufficiently explored.

Objectives

This study aimed to examine the association between ceramide scores and both CVH and ASCVD risk among African-American adults, and to assess differences in ceramide profiles between African-American and White adults.

Materials and methods

A secondary analysis was performed using data from 2 U.S. studies in Minnesota: The Fostering African-American Improvement in Total Health! (FAITH!) Heart Health+[7] study, a community-based study of Black/African-American adults, and the Prevalence of Asymptomatic Left Ventricular Dysfunction (PAVD) study,[6] a community population-based cohort of predominantly White adults. Ethical approval for both studies was granted by the Mayo Clinic Institutional Review Board (IRB No. 21-011103 (Heart Health+) and 15-005100 (PAVD)). All participants provided informed consent before enrollment. Data included demographic characteristics, behavioral factors (e.g., dietary patterns) and clinical measures (e.g., ceramide 16:0, 18:0, 24:0, and 24:1 levels). Ceramide ratios 16:0/24:0, 18:0/24:0 and 24:1/24:0 were obtained. Ceramide scores were calculated on a 12-point scale and categorized into risk levels: low (0–2), intermediate (3–6), moderate (7–9), and high (10–12). The ASCVD risk was assessed using the American College of Cardiology/American Heart Association (ACC/AHA) 10-year pooled cohort risk scores, while CVH was calculated using AHA Life's Essential 8 (LE8) scores.

Continuous variables, including age and mean ASCVD risk scores, were compared between the 2 study cohorts

using the Kruskal–Wallis test. Pearson's χ^2 goodness of fit test was used to compare categorical variables, including categorized ASCVD risk score, sex and the prevalence of cardiovascular risk factors. Within the African-American cohort, Pearson's correlation was used to assess associations between ceramide levels and both LE8 scores and ASCVD risk. Ceramide scores were compared between study groups using multivariable linear regression with robust standard errors (SEs), adjusting for age, sex and cardiovascular risk factors (e.g., hypertension, diabetes, hyperlipidemia, obesity, and tobacco use/cigarette smoking).

All statistical test assumptions were assessed. Continuous variables were assessed for normality using histograms when comparing mean values between White and African-American cohorts. Linearity between variables was examined for correlation analyses. The linear regression model was assessed for linearity, multicollinearity, homoscedasticity, and normality of residuals (see Supplementary data). Robust SEs were used in the model to provide robust estimates of the mean difference in ceramide scores between cohorts in the context of potential heteroscedasticity and non-normality of error residuals. Statistical significance was defined as a $p < 0.05$. All analyses were conducted using R v. 4.2.2 (R Foundation for Statistical Computing, Vienna, Austria).

Results

Fifty-eight African-American adults (mean age (standard deviation (SD): 54.6 (11.9) years, 67.2% women) and 1,103 White adults (mean age (SD): 64.5 (11.9) years, 52.1% women) were included (Table 1). The prevalence of cardiometabolic risk factors was significantly higher among African-American adults than White participants: obesity (67.9% vs 31.9%; $p < 0.001$), hypertension (69.6% vs 41.2%; $p < 0.001$), diabetes (30.4% vs 11.1%; $p < 0.001$), and hyperlipidemia (42.9% vs 29.3%; $p = 0.03$). The mean ASCVD risk scores were comparable between African-American (12.8% (SD = 10.4)) and White individuals (12.6% (SD = 11.1)) ($p = 0.65$). White individuals had higher ceramide levels for 18:0 (0.10 (0.04) vs 0.08 (0.03) μmol/L) and 24:1 (1.17 (0.33) vs 0.91 (0.31) μmol/L) than African-American adults ($p < 0.001$). In the African-American study group, ceramide scores had a weak, nonsignificant association with LE8 scores ($r = -0.29$, $p = 0.071$) and ASCVD risk scores ($r = -0.10$, $p = 0.503$). White individuals showed higher ceramide scores than African-American adults (mean (SD): 3.48 (2.88) vs 1.90 (1.97), $p = 0.001$) in the adjusted model (Table 2).

Discussion

This study found no significant associations between ceramide profiles and CVH or ASCVD risk among African-Americans. Surprisingly, African-American adults had lower ceramide scores than White adults, despite having more adverse cardiometabolic risk factors. This finding aligns with the limited existing literature. The Baltimore Longitudinal Study of Aging found that African-American adults had lower plasma concentrations of most ceramide and dihydroceramide species compared to White adults.[8] Similarly, Buie et al. reported that African-American individuals with metabolic disease had lower levels of ceramides C16 and C20 compared to White individuals.[3] Additionally, total ceramide levels were higher in African-Americans without metabolic disease relative to those with metabolic disease. These findings suggest distinct expression of sphingolipids in African-Americans.

Clinical validation of ceramide scores has been conducted in White populations in the USA and Europe[5,6]; thus, these findings may not be generalizable to African-Americans due to potential differences in genetic and lifestyle factors between these groups. Notably, African-Americans in this study exhibited lower LE8 scores for diet and body mass index (BMI), which, based on existing evidence on the association between adherence to a healthy diet and circulating ceramide levels, should have resulted in higher ceramide levels.[9,10] One explanation could be that ceramide levels are expressed differently in African-Americans than in White individuals. The postulation of genetic differences in sphingolipid expression and blood levels is supported by evidence of higher occurrences of mutations that lower other lipids such as LDL-C among African-Americans.[11] Further genetic studies are warranted to determine whether African-Americans express ceramides differently.

These findings indicate that ceramide scores may have limited utility in assessing cardiovascular risk in African-American adults, potentially leading to suboptimal patient management and exacerbation of existing CVD disparities in this population. Thus, alternative sphingolipid biomarkers or ceramide score cutoffs should be explored. Buie et al.[3] found that sphingosine-1-phosphate (Sph-1P) levels were notably higher in African-Americans than in White and African-American adults without metabolic disease. Given the observed inverse relationship between ceramide levels and cardiometabolic profiles in African-Americans, future research should investigate whether Sph-1P may serve as a more accurate biomarker of CVD risk in this population. Notably, Mielke et al. found that age-related elevations in ceramides were less pronounced among African-Americans. The observation that African-American adults in our study exhibited ASCVD risk levels comparable to those of the White cohort, despite being younger, underscores the need for further validation and investigation of ceramides as CVD biomarkers in African-American populations.

Limitations of the study

While statistically significant differences in ceramide scores were observed between African-American and White participants, a notable limitation is the relatively

Table 1. Cardiovascular risk factors, cardiovascular health and ceramide levels in African-American and White participants

Variable	African-American (n = 58)	White (n = 1,103)	Test, df, p-value
Female sex, n (%)	39 (67.2)	575 (52.1)	$\chi^2 = 5.05$, df = 1, p = 0.025[$]
Age, mean (SD) [years]	54.6 (11.9)	64.5 (9.5)	H (1, n = 1159) = 34.40, p < 0.001[@]
Cardiovascular risk factors, n (%)			
Obesity	38 (67.9)	352 (31.9)	$\chi^2 = 30.84$, df = 1, p < 0.001[$]
Tobacco use/cigarette smoking	17 (30.9)	510 (48.5)	$\chi^2 = 8.13$, df = 2, p = 0.017[$]
Hypertension	39 (69.6)	454 (41.2)	$\chi^2 = 17.69$, df = 1, p < 0.001[$]
Type 2 diabetes	17 (30.4)	106 (11.1)	$\chi^2 = 18.32$, df = 1, p < 0.001[$]
Hyperlipidemia	24 (42.9)	305 (29.3)	$\chi^2 = 4.68$, df = 1, p = 0.031[$]
ASCVD risk, mean (SD)			
Overall ASCVD risk score	12.8% (10.4)	12.6% (11.1)	H (1, n = 886) = 0.211, p = 0.646[$]
ASCVD category*, n (%)			$\chi^2 = 1.99$, df = 3, p = 0.575[$]
Low	12 (23.5)	248 (29.7)	–
Borderline	10 (19.6)	112 (13.4)	
Intermediate	18 (35.3)	302 (36.2)	
High	11 (21.6)	173 (20.7)	
Cardiovascular health[#], mean (SD)			
Total LE8 score	57.4 (11.5)	–	–
LE8 components scores			
Diet	30.3 (12.4)	–	–
Physical activity	50.2 (38.6)	–	–
Tobacco or nicotine exposure	84.1 (24.2)	–	–
Sleep health	73.7 (29.0)	–	–
Body mass index (BMI)	36.4 (30.4)	–	–
Blood lipids (non-HDL-C)	64.6 (29.5)	–	–
Blood glucose	77.1 (29.8)	–	–
Blood pressure	39.1 (32.2)	–	–
Ceramides, mean (SD)			
Ceramide (16:0) [µmol/L]	0.28 (0.06)	0.27 (0.06)	H (1, n = 1150) = 0.36, p = 0.547[@]
Ceramide (18:0) [µmol/L]	0.08 (0.03)	0.10 (0.04)	H (1, n = 1137) = 24.31, p < 0.001[@]
Ceramide (24:1) [µmol/L]	0.91 (0.31)	1.17 (0.33)	H (1, n = 1146) = 33.68, p < 0.001[@]
Ceramide ratio (16:0)/(24:0)	0.08 (0.01)	0.08 (0.02)	H (1, n = 1150) = 0.02, p = 0.889[@]
Ceramide ratio (18:0)/(24:0)	0.02 (0.01)	0.03 (0.01)	H (1, n = 1137) = 21.21, p < 0.001[@]
Ceramide ratio (24:1)/(24:0)	0.25 (0.06)	0.33 (0.11)	H (1, n = 1145) = 44.51, p < 0.001[@]
Ceramide score (0–12)[&]	1.90 (1.97)	3.48 (2.88)	p = 0.001
Ceramide score category, n (%)			$X^2 = 21.38$, df = 3, p < 0.001[$]
0–2	38 (76.0)	475 (43.2)	–
3–6	10 (20.0)	441 (40.1)	
7–9	2 (4.0)	150 (13.6)	
10–12	0 (0.0)	34 (3.1)	

*Risk categories: low (<5%), borderline (5%–7.5%), intermediate (7.5–20%), high (≥20%); [#]Score range: 0–100; categories: low (<50), moderate (50–79), high (≥80); [&]Adjusted for age, sex, hypertension, diabetes, hyperlipidemia, obesity, and smoking status (see Table 2 for full model results); ASCVD – atherosclerotic cardiovascular disease; LE8 – Life's Essential 8; [@]H: Kruskal–Wallis test; df – degrees of freedom; SD – standard deviation; HDL-C – high-density lipoprotein cholesterol; BMI – body mass index.

Table 2. Multivariable linear regression analysis of ceramide risk score with robust SEs

Variable	Estimate	SE	T value	p-value
Intercept	–0.866	0.663	–1.307	0.192
Race (African-American)	–1.189	0.354	–3.361	0.001[#]
Age	0.056	0.010	5.439	<0.001
Sex (female)	0.608	0.191	3.190	0.001
Hypertension (yes)	0.456	0.197	2.309	0.021
Diabetes (yes)	–0.135	0.289	–0.468	0.640
Hyperlipidemia (yes)	–0.642	0.201	–3.190	0.001
Obesity (yes)	0.785	0.205	3.823	<0.001
Smoker (former)	0.073	0.196	0.374	0.708
Smoker (current)	0.579	0.385	1.502	0.133

[*]Model adjusted for age, sex, hypertension, diabetes, hyperlipidemia, obesity, and tobacco use/cigarette smoking status; [#]The p-value for the ceramide score comparison reported in Table 1. SE – standard error.

small sample of African-American participants. This may have limited the ability to detect associations between ceramide profiles and CVH or ASCVD risk within this group, raising the possibility of type II error and the potential for missed meaningful associations. Larger, adequately powered studies are warranted to validate ceramide scores as CVD risk biomarkers or to identify alternative sphingolipid biomarkers to assess CVD risk among African-Americans. Additionally, due to differences in study design, behavioral metrics (diet, physical activity and sleep health) were not collected in the PAVD study,[6] precluding an assessment of LE8 scores among White participants. Given these limitations, this heterogeneity should be considered when interpreting the comparative findings, and our results should be interpreted cautiously as exploratory.

Conclusions

While ceramide profiles are beneficial in predicting CVD risk in White populations, our study findings suggest that their utility in assessing risk among African-American adults appears limited. Identifying accurate population-specific biomarkers or ceramide score cutoffs is crucial for preventing CVD and improving health outcomes in high-risk populations such as African-Americans.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.15756695. The package contains the following files:

Supplementary Fig. 1. Scatterplot to assess linear relationship between LE8 score and ceramide score in the cohort of African-American adults.

Supplementary Fig. 2. Scatterplot to assess linear relationship between ASCVD score and ceramide score in the cohort of African-Americans adults.

Supplementary Fig. 3. Histogram to check normality assumption for age in the cohorts of African-American and White adults.

Supplementary Fig. 4. Histogram to check normality assumption for ASCVD risk score in the cohorts of African-American and White adults.

Supplementary Fig. 5. Histogram to check normality assumption for ceramide 16_0 in the cohorts of African-American and White adults.

Supplementary Fig. 6. Histogram to check normality assumption for ceramide 18_0 in the cohorts of African-American and White adults.

Supplementary Fig. 7. Histogram to check normality assumption for ceramide 24_1 in the cohorts of African-American and White adults.

Supplementary Fig. 8. Histogram to check normality assumption for ceramide ratio (16_0)_(24_0) in the cohorts of African-American and White adults.

Supplementary Fig. 9. Histogram to check normality assumption for ceramide ratio (18_0)_(24_0) in the cohorts of African-American and White adults.

Supplementary Fig. 10. Histogram to check normality assumption for ceramide ratio (24_1)_(24_0) in the cohorts of African-American and White adults.

Supplementary Fig. 11. Scatterplot to assess the linear relationship between age and ceramide score in the cohorts of African-American and White adults.

Supplementary Fig. 12. Plot of model residuals compared to fitted values to assess the validity of the linear regression model in the cohorts of African-American and White adults.

Supplementary Fig. 13. Scale-location plot to assess homoscedasticity in the linear regression model in the cohorts of African-American and White adults.

Supplementary Fig. 14. Q-Q plot to assess normality of residuals in the linear regression model in the cohorts of African-American and White adults.

Supplementary Table 1. Table of generalized variance inflation factor values to assess for multicollinearity in the cohorts of African-American and White adults.

Data Availability Statement

The datasets supporting the findings of the current study are available from the corresponding author on reasonable request. The data are not publicly available due to an agreement with the FAITH! Heart Health+ study participants.

Use of AI and AI-assisted technologies

Not applicable.

Consent for publication

Not applicable.

ORCID iDs

Mathias Lalika https://orcid.org/0000-0003-0229-9942
Vlad C. Vasile https://orcid.org/0000-0001-5447-8091
Matthew P. Johnson https://orcid.org/0000-0002-6341-7220
Sharonne N. Hayes https://orcid.org/0000-0003-3129-362X
Lisa A. Cooper https://orcid.org/0000-0001-6707-6390
Christi A. Patten https://orcid.org/0000-0002-7194-8160
LaPrincess C. Brewer https://orcid.org/0000-0002-6468-9324

References

1. Martin SS, Aday AW, Allen NB, et al. 2025 Heart Disease and Stroke Statistics: A Report of US and Global Data From the American Heart Association. *Circulation*. 2025;151(8):e41–e660. doi:10.1161/CIR.0000000000001303
2. Virani SS, Brautbar A, Davis BC, et al. Associations between lipoprotein(a) levels and cardiovascular outcomes in black and white subjects: The Atherosclerosis Risk in Communities (ARIC) Study. *Circulation*. 2012;125(2):241–249. doi:10.1161/CIRCULATIONAHA.111.045120
3. Buie JNJ, Hammad SM, Nietert PJ, et al. Differences in plasma levels of long chain and very long chain ceramides between African Americans and whites: An observational study. *PLoS One*. 2019;14(5):e0216213. doi:10.1371/journal.pone.0216213
4. Hilvo M, Vasile VC, Donato LJ, Hurme R, Laaksonen R. Ceramides and ceramide scores: Clinical applications for cardiometabolic risk stratification. *Front Endocrinol (Lausanne)*. 2020;11:570628. doi:10.3389/fendo.2020.570628
5. Leiherer A, Muendlein A, Saely CH, Laaksonen R, Fraunberger P, Drexel H. Ceramides improve cardiovascular risk prediction beyond low–density lipoprotein cholesterol. *Eur Heart J Open*. 2023;4(1):oeae001. doi:10.1093/ehjopen/oeae001
6. Vasile VC, Meeusen JW, Medina Inojosa JR, et al. Ceramide scores predict cardiovascular risk in the community. *Arterioscler Thromb Vasc Biol*. 2021;41(4):1558–1569. doi:10.1161/ATVBAHA.120.315530
7. Lalika M, McCoy CR, Jones C, et al. Rationale, design, and participant characteristics of the FAITH! Heart Health+ study: An exploration of the influence of the social determinants of health, stress, and structural racism on African American cardiovascular health. *Contemp Clin Trials*. 2024;143:107600. doi:10.1016/j.cct.2024.107600
8. Mielke MM, Bandaru VVR, Han D, et al. Demographic and clinical variables affecting mid- to late-life trajectories of plasma ceramide and dihydroceramide species. *Aging Cell*. 2015;14(6):1014–1023. doi:10.1111/acel.12369
9. Lankinen M, Schwab U, Kolehmainen M, et al. A healthy Nordic diet alters the plasma lipidomic profile in adults with features of metabolic syndrome in a multicenter randomized dietary intervention. *J Nutr*. 2016;146(4):662–672. doi:10.3945/jn.115.220459
10. Drazba MA, Holásková I, Sahyoun NR, Ventura Marra M. Associations of adiposity and diet quality with serum ceramides in middle-aged adults with cardiovascular risk factors. *J Clin Med*. 2019;8(4):527. doi:10.3390/jcm8040527
11. Cohen J, Pertsemlidis A, Kotowski IK, Graham R, Garcia CK, Hobbs HH. Low LDL cholesterol in individuals of African descent resulting from frequent nonsense mutations in PCSK9. *Nat Genet*. 2005;37(2):161–165. doi:10.1038/ng1509