ISSN 0104-4230
ISSN 1806-9282 (On-line)

# RAMB

## Journal of The Brazilian Medical Association

Volume 72, Number 9
September, 2026

# RAMB

Journal of The Brazilian Medical Association

Volume 72, Number 9, September, 2026

ISSN 0104-4230
ISSN 1806-9282 (On-line)


## SECTIONS


### EDITORIAL

e20260777 The patient statute (Law No. 15,378/2026) and the redefinition of medical authority: autonomy, dignity, and legal certainty at the bedside

### LETTERS TO THE EDITOR

e20251895 Heart rate variability, sleep quality, and quality of life in epilepsy patients are multidependent and cannot be reduced to the quality of seizure control

### SHORT COMMUNICATIONS

e20260147 Quick communication: Active surveillance of rat sarcoma viral oncogene homolog-positive indeterminate thyroid nodules with categories III and IV, the Bethesda System for Reporting Thyroid Cytopathology, in thyroidology for thyroidologists


## ARTICLES


### ORIGINAL ARTICLES

e20250590 Translation, cultural adaptation, and validation of the Genital Self-Image Scale (GSIS-20) for the Portuguese language

e20251444 Psychometric properties of the Brazilian 14-Item Health Literacy Scale for people living with mental disorders: a cross-sectional study

e20252134 Sarcopenia, its components and fracture risk in people living with human immunodeficiency virus

e20260194 Determining women's awareness levels of gynecological cancer and their perceptions of breast cancer fatalism

e20260286 Evaluation of serum soluble triggering receptor expressed on myeloid cells-1 in polycystic ovary syndrome

e20260287 Artificial intelligence-driven triage in emergency departments: a comparative study of machine learning approaches

e20260298 Association between triglyceride–glucose and fibrosis-4 indices and nephrolithiasis: a cross-sectional case–control study

e20260317 Postural imbalance in patients with rheumatoid arthritis: a cross-sectional study

e20260435 Investigation of the relationship between expression patterns of plasma microRNAs and pregnancy complications

e20260456 Functional shoulder performance and fatigue in young adults with forward head posture with or without scapular dyskinesis

e20260490 Thyroid function and anemia: results from the Brazilian Longitudinal Study of Adult Health

e20260500 Self-efficacy in remitted bipolar disorder: associations with employment and self-esteem

e20260564 A comparative analysis of semiannual versus annual rituximab maintenance in patients with refractory idiopathic inflammatory myopathies: a retrospective longitudinal real-world study

e20260575 Inpatient rehabilitation outcomes in acquired brain injury: a retrospective study

e20260577 Cancer-related deaths among cases referred for forensic evaluation in Türkiye: a 5-year retrospective analysis

e20260588 Evaluation of infants aged 0–12 months admitted to the pediatric emergency department: the example of Batman Training And Research Hospital

e20260594 The role of hematological parameters and systemic immune-inflammation index in the management of symptomatic gestational hydronephrosis

e20260604 Factors associated with unsatisfactory ratings among medical courses in the 2025 Brazilian National Medical Education Assessment Exam: a cross-sectional study

e20260646 Osteopathic manual therapy in fibromyalgia: a pilot study

e20260662 Post-cesarean analgesia: transversalis fascia plane block versus wound infiltration in a randomized trial

e20260706 Retrospective study of inflammatory markers in high-grade cervical lesions

e20260753 Artificial intelligence in acne vulgaris patients

e20260760 Lipedema research (2000–2025): a global bibliometric analysis with evidence of an emerging immunological paradigm

e20260783 Night eating syndrome in women with gestational diabetes mellitus: a comparative cross-sectional study

e20260803 Rational Drug Use Support System: design, artificial intelligence integration, and pilot feasibility trial

e20260842 Association of Vascular Endothelial Growth Factor A-2549 I/D polymorphism with polycystic ovary syndrome susceptibility in Turkish women

e20260845 Weight loss–independent associations of body composition changes with device-derived metabolic age during Mediterranean diet follow-up

e20260862 Fundus camera and artificial intelligence for diabetic retinopathy screening in a São Paulo shelter

e20261603 Career, Autonomy, Rights and Equity among Doctors: Brazilian Medical Association survey

e20266403 Pain, anxiety, and comfort among women undergoing medical termination


## ERRATUM


e20251775-ER ERRATUM: Clinical profile of acute myocardial infarction: the Registry of Acute Medical Emergencies registry

RAMB
Journal of The Brazilian Medical Association

## EDITORIAL BOARD

### EDITORS-IN-CHIEF

Renato Delásciо Lopes
Roseli Nomura
José Maria Soares Jr.

### MANAGING EDITOR

Cesar Teixeira

## ASSOCIATED EDITORS

Albert Bousso
Ana Gabriel P. Santos
Ana Pontes
Anna Andrei
Auro Del Giglio
Claudia Leite
Dani Ejzenberg
Dimas Ikeoki
Edna Frasson de S. Montero
Eduardo F. Borba
Edward Araújo Jr
Gabriel Costa Osanan
Isabel Sorpreso
Isabela Giuliano
Lilian Sadeck
Linamara Batistella
Lucia Pellanda
Paulo Kassab
Rachel Riera
Sergio C. Nahas
Werther B. W. de Carvalho

## INTERNATIONAL EDITORS

Frida Leonetti
Geltrude Mingrone
Giuseppe Barbaro
Marcelo Marotti
Walter Ageno

## JUNIOR EDITOR

André Zimerman

## SPECIALTY EDITORS

### ACUPUNCTURE

Sidney Brandão

### ALLERGY AND IMMUNOLOGY

Dirceu Solé

### ANAESTHESIOLOGY

Plínio da Cunha Leal

### ANGIOLOGY AND VASCULAR SURGERY

Edwaldo Edner Joviliano

### CARDIOLOGY

Weimar Kunz Sebba B. de Souza

### CARDIOVASCULAR

Marcela da Cunha Sales

### CLINICAL ONCOLOGY

Alexandre Palladino

### CLINICAL PATHOLOGY / LABORATORIAL MEDICINE

André Doi

### COLOPROCTOLOGY

Henrique Sarubbi Fillmann

### DERMATOLOGY

Flávia Vasques Bittencourt

### DIGESTIVE ENDOSCOPY

Fauze Maluf Filho

### DIGESTIVE SURGERY

Fernando Antônio Siqueira Pinheiro

### EMERGENCY MEDICINE

Hélio Penna Guimarães

### ENDOCRINOLOGY AND METABOLISM

Paulo Augusto Carvalho de Miranda

### FAMILY AND COMMUNITY MEDICINE

Leonardo Cançado Monteiro Savassi

### GASTROENTEROLOGY

Frederico Passos Marinho

### GENERAL SURGERY

Luiz Carlos Von Bahten

### GERIATRICS AND GERONTOLOGY

Hercilio Hoepfner Junior

### GYNAECOLOGY AND OBSTETRICS

Agnaldo Lopes da Silva Filho

### HAND SURGERY

Antônio Tufi Neder Filho

### HEAD AND NECK SURGERY

Leandro Luongo Matos

### HEMATOLOGY AND HEMOTHERAPY

Fernando Ferreira Costa

### HOMEOPATHY

Flavio Dantas de Oliveira

### INFECTIOUS DISEASES

Alexandre Vargas Schwarzbold

INTENSIVE MEDICINE
Israel Silva Maia

INTERNAL MEDICINE
Ana Paula de Oliveira Ramos

LEGAL MEDICINE AND MEDICAL EXAMINATIONS
Rosa Amélia Andrade Dantas

MASTOLOGY
Gil Facina

MEDICAL GENETICS
Ida Vanessa D. Schwartz

NEUROSURGERY
Manoel Jacobsen Teixeira

NEPHROLOGY
Lúcio Roberto Requião Moura

NEUROLOGY
Marcondes Cavalcante França Jr.

NUCLEAR MEDICINE
Diego Pianta

NUTROLOGY
Aline Zanetta

OCCUPATIONAL MEDICINE
Andrea Franco Amoras Magalhães

OPHTHALMOLOGY
Eduardo Melani Rocha

ORTHOPAEDICS AND TRAUMATOLOGY
Sergio Luiz Checchia

OTOLARYNGOLOGY
Thiago Freire Pinto Bezerra

PAEDIATRIC
Lilian dos Santos Rodrigues Sadeck

PAEDIATRIC SURGERY
Lisieux Eyer Jesus

PATHOLOGY
Monique Freire Santana

PHYSICAL MEDICINE AND REHABILITATION
Eduardo de Melo Carvalho Rocha

PLASTIC SURGERY
Daniela Francescato Veiga

PREVENTIVE MEDICINE AND HEALTH ADMINISTRATION
Antônio Eduardo Fernandes D´Aguiar

PSYCHIATRY
Leonardo Rodrigo Baldaçara

PULMONOLOGY / PHTHISIOLOGY
Suzana Erico Tanni Minamoto

RADIOTHERAPY
Wilson José Almeida Jr.

RADIOLOGY
Alexandre Bezerra

RHEUMATOLOGY
Ricardo Machado Xavier

SPORTS MEDICINE
Neuza Mitsuanga

SURGICAL ONCOLOGY
Héber Salvador de Castro Ribeiro

TRAFFIC MEDICINE
José Heverardo da Costa Montal

THORACIC SURGERY
Juliana Dias Nascimento Ferreira

UROLOGY
Roni de Carvalho Fernandes

## ASSOCIAÇÃO MÉDICA BRASILEIRA
(BRAZILIAN MEDICAL ASSOCIATION)

### MANAGEMENT BOARD 2024-2026

PRESIDENT
César Eduardo Fernandes (SP)

1ST VICE-PRESIDENT
Luciana Rodrigues da Silva (BA)

2ND VICE-PRESIDENT
Nerlan Carvalho (PR)

VICE-PRESIDENTS
Etelvino Trindade (DF) – Mid-West
Claudia Navarro Lemos (MG) – Southeast
Paulo Toscano (PA) – North
Bento Bezerra Neto (PE) – Northeast
Juarez Molinari (RS) – South

GENERAL SECRETARY
Florisval Meinão (SP)

1ST SECRETARY
Maria Rita de Souza Mesquita (SP)

1ST TREASURER
Lacildes Rovella Júnior (SP)

2ND TREASURER
Fernando Tallo (SP)

ADMINISTRATIVE DIRECTOR
Akira Ishida (SP)

CULTURAL DIRECTOR
Romulo Capello (RJ)

DIRECTOR OF CORPORATE RELATIONS
Carlos Henrique Mascarenhas (MG)

DIRECTOR OF INTERNATIONAL RELATIONS
Carlos Vicente Serrano Junior (SP)

SCIENTIFIC DIRECTOR
José Eduardo Lutaif Dolci (SP)

ACADEMIC DIRECTOR
Clóvis Francisco Constantino (SP)

DIRECTOR OF MEMBER SUPPORT SERVICES
José Aurillo Rocha (CE)

DIRECTOR OF PARLIAMENTARY AFFAIRS
Luciano Gonçalves de Souza Carvalho (DF)

COMMUNICATIONS DIRECTOR
Luiz Carlos Von Bahten (PR)



## RAMB - REVISTA DA ASSOCIAÇÃO MÉDICA BRASILEIRA
(JOURNAL OF THE BRAZILIAN MEDICAL ASSOCIATION)

**Editors-in-Chief:** Renato Deláscio Lopes, José Maria Soares Jr and Roseli Nomura.
**Managing Editor:** César Teixeira
**E-mail:** ramb@amb.org.br
**Website:** www.ramb.org.br

**ADDRESS:** Rua São Carlos do Pinhal, 324
Bela Vista – São Paulo
Postal Code: 01333-903
Phone no.: (+55 11) 3178-6800 Ext. 177
The RAMB, Journal of The Brazilian Medical Association, is an official publication of the Associacao Medica Brasileira (AMB – Brazilian Medical Association), indexed in Medline, Science Citation Index Expanded, Journal Citation Reports, Index Copernicus, Lilacs, and Qualis A3 Capes databases, and licensed by Creative CommonsR.
Registered in the 1st Office of Registration of Deeds and Documents of Sao Paulo under n. 1.083, Book B, n. 2.

Publication norms are available on the website www.ramb.org.br



All rights reserved and protected by Law n. 9.610 – 2/19/1998. No part of this publication may be reproduced without prior written authorization of the AMB, whatever the means employed: electronic, mechanical, photocopying, recording or other.



**THE RAMB IS INDEXED IN SCIELO - SCIENTIFIC ELECTRONIC LIBRARY ONLINE.**

**Editorial Production**

ZEPPELINI PUBLISHERS



The advertisements and opinions published in the Ramb are the sole responsibility of the advertisers and authors.
The AMB and Zeppelini Publishers are not responsible for its content.

https://doi.org/10.1590/1806-9282.20260777

# The patient statute (Law No. 15,378/2026) and the redefinition of medical authority: autonomy, dignity, and legal certainty at the bedside

Fábio Roberto Cabar[1]*, Bruno Alves Gaia[2]

The 21st century has consolidated an irreversible paradigmatic shift in the physician–patient relationship: the definitive decline of traditional Hippocratic paternalism in favor of strengthening the principle of autonomy[1]. In the national context, this historical evolution reaches its institutional and normative peak with the enactment of Law No. 15,378 of April 6, 2026, which establishes the Patient Statute in Brazil[2]. This legislative milestone assumes an extremely relevant role in the Brazilian healthcare ecosystem, as it transposes purely ethical and deontological guidelines into the realm of cogent legal obligations[2]. The manifest objective of this legislation is to resolve complex bioethical dilemmas and provide much-needed legal certainty to daily clinical practice, with special emphasis on end-of-life scenarios, where the preservation of human dignity must imperatively prevail over the mere artificial biological maintenance of organs and functions[2].

Autonomy, understood under the bioethical aegis as the inalienable right to self-government and self-determination regarding one's own health, stands as the cornerstone of person-centered medicine[1]. However, the transition of the physician's historical role, from an "absolute guardian of life" to a "facilitator of autonomy," imposes a profound and latent moral tension at the bedside[1]. The apparent conflict between the duty of beneficence (acting for the patient's clinical well-being) and respect for the sovereignty of individual will becomes the epicenter of complex debates within healthcare institutions[1]. Although the Brazilian Code of Medical Ethics had already signaled this horizontalization in recent decades, the positive law brought by the 2026 Statute alters the level of the discussion, transforming conduct recommendations into binding normative mandates[3].

This dilemma manifests with dramatic contours and greater intensity when the strictly technical convictions of the care team collide with the sovereign decision of the patient to refuse a specific treatment[3]. From this scenario emerges the critical concept of moral distress: the psychological and existential exhaustion experienced by the healthcare professional when compelled by force of law to follow a course of action that directly contradicts their clinical and technical perception of what would be the "best interest" of the patient[3]. It is imperative to understand, however, that in contemporary bioethics, beneficence cannot be exercised in an impositive or authoritarian manner; it must obligatorily be filtered and reframed through the prism of the values, beliefs, and biography of the individual experiencing the illness process[4].

Faced with such impasses, the 2026 Patient Statute offers fundamental practical and operational dimensions for the peaceful resolution of conflicts within the hospital environment[2]. First, the new legislation grants strictly binding force to Advance Healthcare Directives, ensuring unequivocally that the individual's so-called "living will" is fully respected by the medical team and administrators, even in situations of supervening cognitive incapacity[5]. This legal determination acts as a powerful reducer of ethical clashes and exhausting litigation between the clinical body and the patient's family members[5].

Furthermore, the legal text provides robust and explicit legal protection for the practice of orthothanasia[6]. Consequently, physicians who decide to suspend or abstain from introducing futile, invasive, or extraordinary treatments, characterized as therapeutic obstinacy or dysthanasia, are legally safeguarded, redirecting the assistance focus toward comprehensive palliative care focused on relieving physical, psychological, and spiritual

[1]Universidade de São Paulo, School of Medicine – São Paulo (SP), Brazil.
[2]Universidade Paulista – Campus Alphaville, Instituto de Ciências da Saúde – São Paulo (SP), Brazil.
*Corresponding author: fabio.cabar@hc.fm.usp.br
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on May 18, 2026. Accepted on June 14, 2026.
Scientific Editor: Roseli Nomura

suffering[6]. Such normative advancements culminate in greater predictability and legal protection for professional practice, defining clear boundaries for medical liability in the face of a patient's informed refusal, and mitigating the deep-seated fear of judicial litigation that historically impelled professionals toward the practice of defensive and tortuous medicine[2].

The contemporary challenge posed to the medical community requires professionals to make a definitive postural transition: to abandon the traditional purely prescriptive and directive stance and adopt a genuinely dialogical posture, validating and welcoming the prospective autonomy of the patient[5]. Ultimately, the Brazilian Patient Statute does not diminish, weaken, or subvert the scientific and moral authority of the physician; on the contrary, it recontextualizes it and elevates it to a new level of humanistic sophistication[2]. Recognizing the biological limits of therapeutics, accepting the finitude inherent to the human condition, and allowing death to occur in its natural time constitute the highest and most noble expressions of comprehensive medical care[6]. The true ethical commitment of 21st-century medicine lies in the firm conviction that when a cure is no longer a technically achievable horizon, dignity must remain the supreme value guiding both the law and the art of caring[6].

## AUTHORS' CONTRIBUTIONS

**FRC:** Conceptualization, Formal Analysis, Writing – original draft. **BAG:** Conceptualization, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Beauchamp TL, Childress JF. Principles of biomedical ethics. 8th ed. New York: Oxford University Press; 2019.
2. Brasil. Lei nº 15.378, de 6 de abril de 2026. Institui o estatuto dos direitos do paciente. Diário Oficial da União; 2026. Seção 1:1.
3. Conselho Federal de Medicina. Código de ética médica: resolução CFM nº 2.217/18. Brasília (DF): CFM; 2019.
4. Tabet L, Garrafa V. Fim da vida: morte e eutanásia. Rev Bras Bioét. 2016;12:1-16.
5. Dadalto L. Diretivas antecipadas de vontade: um modelo prático. 5th ed. Indaiatuba (SP): Foco; 2020.
6. Schramm FR. Finitude e bioética do fim da vida. Rev Bras Cancerol. 2012;58(1):23-30.

https://doi.org/10.1590/1806-9282.20251895

# Heart rate variability, sleep quality, and quality of life in epilepsy patients are multidependent and cannot be reduced to the quality of seizure control

Sinda Zarrouk[1], Josef Finsterer[2]*

Dear Editor,

We read with interest the article by Duran et al. on a prospective cross-sectional study of the relationship between heart rate variability (HRV), quality of life, and sleep quality in 27 patients with drug-controlled epilepsy and 23 patients with drug-resistant epilepsy[1]. It was found that HRV was reduced in patients with epilepsy and that HRV correlated with cognitive function and well-being[1]. The study is interesting, but some points need to be discussed.

The first point is that subjective assessment of sleep quality using a questionnaire such as the Pittsburgh Sleep Quality Index may not adequately capture sleep quality. Only polysomnography can objectively assess whether sleep architecture is impaired or not. It is therefore recommended that the study be repeated using more sophisticated methods.

The second point is that sleep quality may depend not only on the quality of seizure control, but also on various other influencing factors. Numerous exogenous and endogenous factors also determine sleep quality and sleep hygiene[2]. These additional determinants of sleep quality must be taken into account when assessing the sleep performance of epilepsy patients and interpreting the study results. Adequate and restful sleep is a prerequisite for seizure control.

The third problem is that comorbidities were not taken into account as determinants of HRV, quality of life, and sleep quality[1]. Various conditions, such as myocardial disease, cardiac conduction disorders, alcoholism, obesity, psychiatric disorders, heat, noise, shift work, harmful and dangerous substances, orthopedic disorders, rheumatological disorders, immunological disorders, malignancies, and movement disorders can affect HRV, quality of life, and sleep[3]. Failure to consider comorbidities in the analysis could distort the results or even lead to false negative or false positive results.

The fourth point is that all antiseizure medications (ASMs) are associated with side effects[4]. Phenytoin and carbamazepine can reduce HRV[5]. Carbamazepine and phenobarbital impair sleep architecture[6]. Most ASMs reduce rapid eye movement sleep[6].

The fifth point is that low HRV has been attributed to reduced parasympathetic tone[1], suggesting that autonomic innervation is disrupted in epilepsy patients. Therefore, it is crucial to know how many of the included patients had polyneuropathy with autonomic involvement or pure autonomic neuropathy. Autonomic neuropathy can be caused by seizures that affect the central autonomic network or by the long-term effects of certain ASMs. Autonomic dysfunction may also be due to impairment of the central parasympathetic pathways.

In summary, all determinants of sleep quality, quality of life, and HRV must be considered when assessing their significance for epilepsy. As long as the interdependence of these parameters is not taken into account, their evaluation in epilepsy patients remains unclear.

## AUTHORS' CONTRIBUTIONS

**SZ:** Validation, Writing – review & editing. **JF:** Conceptualization, Formal Analysis, Writing – original draft.

## DATA AVAILABILITY STATEMENT

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

[1]University of Tunis El Manar and Genomic Platform, Pasteur Institute of Tunis – Tunis, Tunisia.
[2]Neurology and Neurophysiology Center, Department of Neurology – Vienna, Austria.
*Corresponding author: fifigs1@yahoo.de
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on October 21, 2025. Accepted on November 09, 2025.
Scientific Editor: Roseli Nomura

## REFERENCES


1. Duran S, Boduroglu Y, Celikbilek A. Heart rate variability, quality of life, and sleep quality in patients with epilepsy. Rev Assoc Med Bras (1992). 2025;71(7):e20250002. https://doi.org/10.1590/1806-9282.20250002
2. Altun I, Cınar N, Dede C. The contributing factors to poor sleep experiences in according to the university students: a cross-sectional study. J Res Med Sci. 2012;17(6):557-61. PMID: 23626634
3. Butterly EW, Hanlon P, Shah ASV, Hannigan LJ, McIntosh E, Lewsey J, et al. Comorbidity and health-related quality of life in people with a chronic medical condition in randomised clinical trials: an individual participant data meta-analysis. PLoS Med. 2023;20(1):e1004154. https://doi.org/10.1371/journal.pmed.1004154
4. Nei M, Ho J, Ho RT. Cardiovascular effects of antiseizure medications for epilepsy. CNS Drugs. 2025;39(4):383-401. https://doi.org/10.1007/s40263-025-01163-x
5. Tomson T, Kennebäck G. Arrhythmia, heart rate variability, and antiepileptic drugs. Epilepsia. 1997;38(11 Suppl):S48-51. https://doi.org/10.1111/j.1528-1157.1997.tb06128.x
6. Carvalho BMS, Chaves J, Silva AM. Effects of antiepileptic drugs on sleep architecture parameters in adults. Sleep Sci. 2022;15(2):224-44. https://doi.org/10.5935/1984-0063.20220045

SHORT COMMUNICATION

https://doi.org/10.1590/1806-9282.20260147

# Quick communication: Active surveillance of rat sarcoma viral oncogene homolog-positive indeterminate thyroid nodules with categories III and IV, the Bethesda System for Reporting Thyroid Cytopathology, in thyroidology for thyroidologists

Ilker Sengul[1,2,3]*, Demet Sengul[4]

Indeterminate cytology remains a major source of controversy in the management of suspicious thyroid nodules in thyroidology today[1-10]. The prospective analysis by Wu et al.[1], offers a timely provocation regarding the clinical management of RAS-mutated indeterminate[1-7] thyroid nodules, just the Categories III and IV, the Bethesda System for Reporting Thyroid Cytopathology (TBSRTC), in thyroidology[1-10]. By demonstrating that immediate resection may lead to overtreatment in up to 50% of cases, the authors provide a necessary evidentiary basis for active surveillance (AS) in thyroidology. However, a multifaceted interrogation of their data—particularly the longitudinal growth dynamics and the nuanced molecular landscape—suggests that our understanding of "Rat sarcoma viral oncogene homolog (RAS)-positivity" requires a more granular, isoform-specific refinement. The most striking, yet perhaps under-scrutinized, finding in Wu et al.[1] cohort is the nonlinear growth trajectory of these nodules. While volume remained remarkably stable during the first 24 months (2% cumulative growth), a sharp escalation to 24% growth was observed by the 4-year mark. This "growth spurt" challenges the assumption of perpetual indolence and suggests a biological "tipping point." One might ask: Is this delayed progression indicative of a late-stage clonal expansion or a manifestation of specific RAS isoforms? As such, evidence increasingly suggests that the clinical "aggressiveness" of a RAS mutation is not monolithic but isoform-dependent. In addition, while Neuroblastoma-rat sarcoma viral oncogene homolog (N-RAS) mutations often correlate with a more indolent, Non-invasive Follicular Thyroid Neoplasm with Papillary-like Nuclear Features (NIFTP)-like phenotype, Harvey-rat sarcoma viral oncogene homolog (H-RAS) and Kirsten-rat sarcoma viral oncogene homolog (K-RAS) isoforms have been associated with a higher propensity for structural progression and, in some contexts, higher-grade histopathology. Furthermore, the observation that nuclear atypia was significantly more prevalent in the surgical group (52 vs. 38%) reinforces the necessity of integrating cytomorphological nuances with molecular data. This is particularly relevant in light of the 2023 TBSRTC, 3rd edition, which emphasizes the subcategorization of Atypia of undetermined significance (AUS) to better predict malignancy risks[2-5]. Moreover, the transition of four patients to surgery—resulting in low-risk malignant histology—may validate the safety of deferred intervention but simultaneously underscores the limitations of binary molecular testing. Herein, without accounting for secondary co-mutations or specific RAS isoforms, we risk applying a "one-size-fits-all" surveillance protocol to a biologically heterogeneous group of lesions. We might liken this situation to an observer who, upon receiving a storm warning (RAS mutation), decides whether to stay outside by examining the structure of the clouds in the sky and the wind speed (clinical factors and ultrasound) instead of immediately taking shelter. Wu et al.[1] attempted to show us with evidence that not every cloud brings a storm. Yet, the 4-year data might serve as a reminder that the horizon remains unpredictable. Intercalary, the authors evaluated indeterminate cytology[1-7] solely by including Categories III and IV, excluding Category V. Future management frameworks must transcend the "positive/negative" dichotomy, moving toward a multi-parametric risk score that integrates isoform specificity, volumetric doubling time, and the refined TBSRTC criteria from the

[1]Giresun University, Faculty of Medicine, Thyroidology Unit – Giresun, Turkey.
[2]Giresun University, Faculty of Medicine, Division of Endocrine Surgery – Giresun, Turkey.
[3]Giresun University, Faculty of Medicine, Department of General Surgery – Giresun, Turkey.
[4]Giresun University, Faculty of Medicine, Department of Pathology – Giresun, Turkey.
*Corresponding author: ilker.sengul.52@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on January 14, 2026. Accepted on January 30, 2026.
Scientific Editor: José Maria Soares Júnior

last edition[2-5]. Only then can we truly fulfill the promise of personalized thyroidology for thyroidologists. *Felix qui potuit rerum cognoscere causas*.

## AUTHORS' CONTRIBUTIONS

**IS:** Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Validation, Visualization, Writing – original draft, Writing – review & editing.

**DS:** Conceptualization, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Wu J, Qui J, Khalil M, Ghaznavi S, Yeo C, Nohr E, et al. Prospective results for active surveillance of RAS-positive indeterminate thyroid nodules with Bethesda III and IV cytologies. Can J Diabetes. 2025;49:S56-62.
2. Sengul I, Sengul D. Blurred lines for management of thyroid nodules in the era of atypia of undetermined significance/follicular lesion of undetermined significance: novel subdivisions of categories IIIA and IIIB in a possible forthcoming The Bethesda System for Reporting Thyroid Cytopathology, 3rd edition; amending versus unnecessary? Rev Assoc Med Bras (1992). 2021;67(10):1385-6. https://doi.org/10.1590/1806-9282.20210763
3. Sengul I, Sengul D. Focusing on thyroid nodules in suspense: 10-15 mm with repeat cytology, Category III, the Bethesda System for Reporting Thyroid Cytopathology, TBSRTC. Rev Assoc Med Bras (1992). 2021;67(2):166-7. https://doi.org/10.1590/1806-9282.67.02.20200828
4. Sengul D, Sengul I. Subdivision of intermediate suspicion, the 2021 K-TIRADS, and category III, indeterminate cytology, the 2017 TBSRTC, 2nd edition, in thyroidology: let bygones be bygones? Ultrasonography. 2023;42(4):600-1. https://doi.org/10.14366/usg.23113
5. Sengul I, Sengul D. Evangely, the subcategorization has been announced in the 2023 Bethesda system for reporting thyroid cytopathology: let bygones be bygones in thyroidology! Rev Assoc Med Bras (1992). 2024;70(3):e20231511. https://doi.org/10.1590/1806-9282.20231511
6. Sengul I, Sengul D. Comment on: "Evaluating treatment options in managing thyroid nodules with indeterminate cytology of TBSRTC in thyroidology: *addendum aut non*?" Rev Assoc Med Bras (1992). 2022;68(7):973-4. https://doi.org/10.1590/1806-9282.20220383
7. Sengul I, Sengul D. May 25-31, International Thyroid Awareness Week & May 25, World Thyroid Day, 2022: indetermination of indeterminate cytology, AUS/FLUS, FN, SUSP, in Thyroidology? Sanamed. 2022;17(2):109-10. https://doi.org/110. 10.5937/sanamed.17-38153
8. Soares Junior JM, Albayrak M, Sengul D, Sengul I. Thyroid function after menopause: is there any concern in thyroidology? Rev Assoc Med Bras (1992). 2024;70(12):e7012EDI. https://doi.org/10.1590/1806-9282.7012EDI
9. Soares Junior JM, Detanac D, Sengul I, Dugalic S, Sengul D, Detanac D. Melatonin, menopause, and thyroid function in gynecologic endocrinology: what is the role? Rev Assoc Med Bras (1992). 2024;70(1):e701EDIT. https://doi.org/10.1590/1806-9282.701EDIT
10. Sengul D, Sengul I, Soares Junior JM. Repercussion of thyroid dysfunctions in thyroidology on the reproductive system: conditio sine qua non? Rev Assoc Med Bras (1992). 2022;68(6):721-2. https://doi.org/10.1590/1806-9282.20220255

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20250590

# Translation, cultural adaptation, and validation of the Genital Self-Image Scale (GSIS-20) for the Portuguese language

Denise Saquetto[1], Renata Almeida de Souza Aranha e Silva[1], Aline Marina Marcondes Braga Rosa[1], Mariane Castiglione[1], Thaiane Suleiman Saoud[1], César Eduardo Fernandes[1], Luiz Gustavo Oliveira Brito[2], Emerson Oliveira[1]*

## SUMMARY

**OBJECTIVE:** The aim of this methodological study was to translate, culturally adapt, and validate the Genital Self-Image Scale (GSIS-20) into Brazilian Portuguese.

**METHODS:** The cross-cultural adaptation followed a standardized five-step process: (1) initial translation into Portuguese by two independent bilingual translators; (2) synthesis of the translations; (3) back-translation into English by two independent native speakers; (4) expert committee review to ensure semantic, idiomatic, experiential, and conceptual equivalence; and (5) pre-test with 15 women to evaluate comprehension and cultural relevance. Subsequently, the psychometric properties of the adapted instrument were assessed in 148 women divided into three groups: nulliparous (n=53), primiparous without pelvic organ prolapse (n=52), and women with stage III or IV pelvic organ prolapse (n=43). The Genital Self-Image Scale (GSIS-20) was applied at three time points to assess internal consistency (Cronbach's alpha), test–retest, and interobserver reliability (intraclass correlation coefficient), and discriminant validity.

**RESULTS:** All subscales of the Genital Self-Image Scale (GSIS-20) showed statistically significant differences between groups. The instrument demonstrated excellent internal consistency (Cronbach's alpha=0.925) and reproducibility (intraclass correlation coefficient >0.75).

**CONCLUSION:** The Genital Self-Image Scale (GSIS-20) was successfully translated and culturally adapted into Brazilian Portuguese, showing satisfactory psychometric properties for use in this population.

**KEYWORDS:** Female genital. Pelvic organ prolapse. Validation study. Translation. Quality of life.

## INTRODUCTION

Female genital self-image is defined as the perception of one's own genitalia, crucial to female sexuality from a biopsychosocial standpoint[1,2]. Positive genital and body self-image improves quality of life. Women with a positive genital self-image exhibit higher self-esteem, sexual confidence, and quality of life while feeling secure in seeking gynecological care. Conversely, negative perceptions can lead to anxiety, depression, sexual dysfunctions, and pursuit of aesthetic procedures[2-7].

Many women increasingly question their genital appearance, often associating a negative vulva image with issues like odor, size, and shape, impacting sexual quality of life[6]. Body and genital image issues may be overlooked during pregnancy, postpartum, or with pelvic dysfunctions like urinary incontinence or pelvic organ prolapse, adversely affecting sexual function[4]. Surgical interventions to address these dysfunctions can enhance sexual satisfaction by improving body and genital image[8].

There is a growing demand for surgical procedures to enhance genital morphology, self-esteem, functional comfort, and sexual pleasure, with cosmetic gynecology expanding rapidly as women seek the "ideal" vulva[3]. Common treatments, albeit not evidence-based, include cosmetics, piercings, and female cosmetic genital surgery[5]. Media influence is significant in these trends, with labiaplasty being prevalent for appearance or comfort concerns[3]. Non-surgical options like laser and radiofrequency are rising, requiring more evidence[9]. These issues call for an interdisciplinary approach considering psychological aspects[10].

The Genital Self-Image Scale (GSIS) was originally developed in the United States by Berman et al. in 2003 to evaluate women's perceptions of their genitalia as an essential component of sexual health[1]. The original version comprised 29 items, which

[1]Centro Universitário Faculdade de Medicina do ABC, Department of Urogynecology and Pelvic Floor Dysfunction, Discipline of Gynecology – Santo André (SP), Brazil.
[2]Universidade de Campinas, School of Medical Sciences, Department of Obstetrics and Gynecology – Campinas (SP), Brazil.
*Corresponding author: emerson.oliveira@fmabc.br
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 04, 2025. Accepted on August 07, 2025.
Scientific Editor: Roseli Nomura

was later shortened and psychometrically validated in a 20-item version (GSIS-20) by Zielinski et al. in 2012[7]. The GSIS-20 assesses four distinct domains: genital confidence, appeal, function, and comfort. Each domain is scored using a four-point Likert scale, with reverse scoring applied to positively phrased items. Despite its validation in Turkey, the United States, and Egypt[4,11,12], there is no culturally adapted version for Brazilian Portuguese. Given that instruments developed in different linguistic and cultural contexts may not be directly applicable elsewhere, cross-cultural adaptation and psychometric validation are essential steps before use in new populations.

## METHODS

### Study design and population

We conducted a cross-sectional study from February to December 2018, approved by the Institutional Review Board of the Centro Universitário FMABC (approval number 2,418,543). The sample included 151 women, divided into three groups: 56 nulliparous (GI), 53 primiparous without Pelvic Organ Prolapse (POP) (GII), and 56 with stage III and IV POP (GIII), assessed by the POP-Q[13]. Participants, aged 25–80, were evaluated by the Urogynecology and Pelvic Floor team at Centro de Atenção Integral à Saúde da Mulher, São Bernardo do Campo, São Paulo, Brazil, an affiliated institution with Faculdade de Medicina do ABC (FMABC).

Inclusion criteria were consent to participate, biological female sex, female gender identity, and literacy. Exclusions were pregnant women and those with prior POP or labiaplasty surgeries.

### Translation and validation of the instrument

The GSIS-20, developed by Berman et al.[1] and validated by Zielinski et al.[7], was used. Authorization was obtained from Professor Laura Berman for translation and validation. The GSIS-20 assesses four domains: genital confidence, appeal, function, and comfort. Scores range from 0 to 40, with higher scores indicating greater satisfaction. Items are on a Likert scale from 0 to 3, and items with positive words are reversed.

The guidelines by Beaton et al.[14] guided the translation process. It included translation into Portuguese, cultural adaptation by a medical team for consensus (>50% agreement), back-translation into English to compare with the original, and a pilot study with 30 patients for feedback on comprehensibility.

### Instrument and procedures

Post-informed consent and medical visit, sociodemographic data were collected. Participants self-administered the GSIS-20 without clarification, completing it at T0 (after the visit), T1 (15 min after T0), and T2 (15 days later), overseen by researchers Denise Saquetto and Aline Marina Marcondes Braga Rosa. Measures included internal consistency, test-retest, and interobserver reliability. Internal consistency assesses item correlation; a retest occurred 15 days after the first assessment. Interobserver reliability was checked for 15 min post-interview, while discriminant validity involved applying the questionnaire to all three groups.

Cronbach's alpha assessed intraclass correlation coefficient (ICC): >0.9 is excellent; 0.7–0.9 is good; 0.6–0.7 is questionable; 0.5–0.6 is poor; <0.5 is unacceptable[15]. When ICC=0, no reproducibility exists; ICC=1 indicates maximum reproducibility; ICC <0.4 is poor; ICC≥0.75 is excellent; and 0.4≤ICC<0.75 is satisfactory[16].

### Statistical analysis

Data were tabulated in Google Drive spreadsheets, analyzed using Prism 8.4.1 and Statistical Package for the Social Sciences. Normality was assessed with the Kolmogorov-Smirnov test. Analysis of Variance (ANOVA), Kruskal-Wallis, and Dunn tested continuous variables, while chi-squared tested categorical variables, with all tests two-tailed and a 5% significance level. Sample size was not calculated due to variability in suggested formulas for translation and validation studies[17]. Reliability was confirmed with ICC for comparisons at T0 and T2 (intraobserver) and T0 and T1 (interobserver)[16].

## RESULTS

We included 53 nulliparous women (GI), 52 primiparous women (GII), and 43 multiparous women with stage III or IV POP (GIII), totaling 148 patients recruited during the study period. We had to exclude 14 women from the study (due to having undergone surgeries for POP correction and/or labiaplasty) during the recruitment period. Demographic data are described in Table 1. Women with POP had a higher mean age than those in GII and GI (64.86±8.97 vs. 37.12±5.54 and 35.04±6.34 years, respectively; p<0.0001), as well as a higher body mass index (BMI) (26.92±3.63 6 kg/m$^2$; p=0.028), a greater proportion of married women (86%; p<0.0001), and higher parity [3 (1–10); p<0.001] than other groups. The rate of hypertension was higher among women with POP 60.6% (p<0.0001) than in the other groups.

Table 2 shows the analysis of the GSIS-20 scores among the three groups. All GSIS-20 subscales showed statistically significant differences between women with POP versus nulliparous women and primiparous women without the disease, confirming that the questionnaire presents discriminant validity.

Internal consistency coefficients (Cronbach's alpha) of the GSIS-20 questionnaire are shown in Table 3. Values above 0.70

**Table 1.** Demographic and clinical baseline characteristics.

| Variable | GI (n=53) | GII (n=52) | GIII (n=43) | p |
|---|---|---|---|---|
| Age[a] | 35.04 (6.34) | 37.12 (5.54) | 64.86 (8.97) | <0.0001 |
| BMI (kg/m²)[a] | 25.05 (4.17) | 26.77 (5.32) | 26.92 (3.63) | 0.028 |
| Ethnicity[b] | | | | |
| White | 41 (77.3) | 37 (71.1) | 26 (60.4) | 0.194 |
| Non-white | 12 (22.4) | 15 (28.9) | 17 (39.6) | |
| Marital status[b] | | | | |
| Married | 22 (41.5) | 35 (67.3) | 37 (86) | <0.0001 |
| Single | 31 (58.5) | 17 (32.7) | 6 (14) | |
| Education level[b] | | | | |
| Primary education | 0 (0) | 0 (0) | 30 (69.8) | <0.0001 |
| Secondary education | 16 (30.2) | 22 (42.3) | 11 (25.6) | |
| Higher education | 37 (69.2) | 30 (57.7) | 2 (4.6) | |
| Parity[c] | 0 | 1 | 3 (1–10) | <0.001 |
| Smoking[b] | 4 (7.5) | 3 (5.8) | 2 (4.6) | 0.8343 |
| Diabetes[b] | 2 (3.8) | 1 (1.9) | 2 (4.6) | 0.5761 |
| Hypertension[b] | 7 (13.2) | 8 (15.4) | 26 (60.6) | <0.0001 |
| Vascular disease[b] | 4 (7.5) | 0 (0) | 2 (4.6) | 0.1423 |

[a]Values presented as the mean and standard deviation. Statistical test: ANOVA. [b]Values presented as numbers and percentages. Statistical test: chi-squared. [c]Values presented as median and variation. Statistical test: Kruskal-Wallis. BMI: body mass index. GI: nulliparous women; GII: primiparous women; GIII: multiparous women with stage III or IV pelvic organ prolapse (POP).

**Table 2.** Discriminant validity of the Genital Self-Image Scale-20 questionnaire between groups.

| Questionnaire domain | GI (n=53) | GII (n=52) | GIII (n=43) | p |
|---|---|---|---|---|
| Genital confidence[a] | 10.34 (1.46) | 10.02 (1.82) | 5.86 (3.80) | <0.0001 |
| Genital appeal[a] | 9.67 (2.53) | 9.80 (2.52) | 4.48 (3.22) | <0.0001 |
| Genital function[a] | 9.28 (1.62) | 9.00 (1.63) | 5.18 (3.67) | <0.0001 |
| Genital comfort[a] | 5.41 (0.74) | 5.29 (0.72) | 4.14 (1.95) | <0.0004 |
| Total score[a] | 34.72 (4.84) | 34.12 (5.10) | 19.67 (10.62) | <0.0001 |

[a]Values presented as the mean and standard deviation. Statistical test: Kruskal-Wallis. Multiple comparisons: GIII<GII and GIII<GI. Statistical test: Dunn's multiple comparison test. GI: nulliparous women; GII: primiparous women; GIII: multiparous women with stage III or IV pelvic organ prolapse (POP).

**Table 3.** Internal consistency coefficients (Cronbach's alpha), test–retest intraobserver and interobserver reliability of the Genital Self-Image Scale (GSIS-20) questionnaire.

| Questionnaire domain | Number of items | Cronbach's alpha | Intraobserver ICC CI | Interobserver ICC CI |
|---|---|---|---|---|
| Genital confidence | 05 | 0.887 | 0.965 (0.951–0.974) p=0.0001 | 0.986 (0.981–0.990) p=0.0001 |
| Genital appeal | 06 | 0.835 | 0.966 (0.953–0.975) p=0.0001 | 0.936 (0.912–0.954) p=0.0001 |
| Genital function | 05 | 0.830 | 0.984 (0.978–0.989) p=0.0001 | 0.936 (0.910–0.954) p=0.0001 |
| Genital comfort | 04 | 0.734 | 0.976 (0.967–0.983) p=0.0001 | 0.939 (0.915–0.956) p=0.0001 |
| Total score | 20 | 0.925 | 0.988 (0.984–0.992) p=0.0001 | 0.969 (0.958–0.977) p=0.0001 |

ICC: intraclass correlation coefficient; CI: confidence interval.

were found for all domains, and for the overall score, Cronbach's alpha coefficient was 0.925, showing excellent internal consistency. Therefore, the reliability of the instrument was adequate.

Finally, no differences were observed in the test–retest comparison of the GSIS-20 translated into Brazilian Portuguese in the intra- or interobserver evaluation. The ICC results have shown excellent reproducibility (ICC>0.75) (Table 3). The final translated version is also available with this manuscript (Supplementary Appendix).

## DISCUSSION

Each society has its own culture, habits, rules, so the translation of a quality-of-life questionnaire should be adapted to the culture of the country. It should have a clear language and be a simple tool that people can answer without difficulty. A questionnaire must be so adapted before the instrument is used in clinical practice or research in populations that speak a different language so that a unified evaluation standard is achieved[18]; that's why a pilot phase was performed so that regionalization was considered in this process. A questionnaire that is valid and reliable in a specific language may not be valid and reliable when used in a different population and may fail to reproduce the same findings when used in other scenarios or by different observers[19].

This study adhered to the methodological framework proposed by Beaton et al.[14] for the cross-cultural adaptation of self-report instruments. By following all five recommended stages—translation, synthesis, back-translation, expert committee review, and pre-testing—the research ensured the semantic, idiomatic, experiential, and conceptual equivalence of the GSIS-20. These methodological safeguards are fundamental to preserving the validity of psychometric instruments when applied to populations culturally and linguistically distinct from those for which the instrument was originally developed.

The original GSIS questionnaire was validated in 2003 in English, and it had 29 items[1]. In 2012, it was shortened and validated with reliability measures for 20 items, again in English[7]. The findings of the present study demonstrate acceptable psychometric properties of the Brazilian Portuguese version of the GSIS-20, which allows its use in this population. The general understanding of the translated terms was adequate, and only small changes were necessary. There was adequate internal consistency for all domains of the questionnaire, and Cronbach's alpha coefficient showed excellent internal consistency of the overall score. We also observed, in the test–retest comparison of the GSIS-20, that the correlations were comparable to those of the validation process of Zielinski et al.[7] and Pakpour et al.[20]. The present study also confirmed the usefulness of the GSIS-20 to assess the impact of POP on the genital image of the participants and thereby demonstrated the discriminant validity. We found that GIII had a score 19.67 points lower than GII and GI ($p<0.0001$).

Our data suggest that dissatisfaction with genital self-image is associated with demographic and clinical characteristics that impair self-perception. Older, less educated, multiparous, and POP women showed greater dissatisfaction with their genitalia. Thus, our results are in line with the study by Zielinski et al., 2012, which showed that women with POP exhibit genital image deterioration[21].

Dissatisfaction with genital image reduces the frequency of visits to the gynecologist and the compliance with preventive exams. Negative body image is associated with worse genital self-image[22,23]. This dissatisfaction may also be caused by the lack of information regarding the anatomy of the genitals. Yulevitch et al., 2013 reported that the lack of knowledge about genitalia and the lack of sexual education may be caused by the cultural context, rigid education, shame, and taboos that still exist in societies[2]. During the data collection process, our team observed that participants lacked basic information about the anatomy of the genitals. In addition, it is common for physicians to feel uncomfortable asking them about these issues. This lack of communication contributes to the poor genital self-image.

Women need to know that each vulva has its own shape, color, size, symmetry, odor, and amount of hair, and each vulva is "normal" and unique, with its own characteristics. A study developed by Sharp and Tiggemann, presented to the participants images of real and healthy female genitals, in all their diversity, and a video showing the digital changes that corrected the protruding small labia in models for pornographic magazines. The authors concluded that the video significantly increased the perceptions of women about the diversity of the appearance of the genitals and that the volunteers who watched it transmitted their knowledge to help other women[24].

This study has strengths and limitations. One limitation is that the pre-test sample included only 15 participants, which falls short of the 30 participants recommended by Beaton et al.[14], potentially limiting the generalizability of the findings. In addition, we did not assess construct validity by comparing the GSIS-20 with other related instruments, which should be addressed in future research. Nonetheless, the GSIS-20 demonstrated excellent internal consistency, reproducibility, and discriminant validity in the Brazilian population, supporting its applicability in both clinical and research contexts.

A recent validation study in Brazil for the Body Image in Pelvic Organ Prolapse Questionnaire did not find association between prolapse staging and genital body image or with sexual function; however, an impaired body image was associated with

worse perception of attractiveness and increased risk for sexual dysfunction[25]. Future studies comparing GSIS with other instruments are needed. Responsiveness was another psychometric variable that was not tested. As for strengths, we have tested most of the psychometrics variables, the presence of two examiners applying the instruments, and the high scores that were found among our data. We believe that the GSIS-20 translated into Portuguese in Brazil is an instrument that can be applied in gynecological visits, contributing to the doctor–patient relationship, since it can evaluate female feelings regarding their genitals and thus help identify factors that contribute to sexual dysfunctions.

## CONCLUSION

The GSIS-20 questionnaire was successfully translated and culturally adapted into Brazilian Portuguese following international guidelines. It demonstrated excellent psychometric properties, including high internal consistency, reproducibility, and discriminant validity. The Brazilian version of the GSIS-20 can be reliably used in clinical and research settings to assess women's genital self-image, particularly in the context of pelvic floor disorders.

## AUTHORS' CONTRIBUTIONS

**DS:** Conceptualization, Investigation, Methodology, Project administration, Writing – original draft, Writing – review & editing. **RASAS:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **AMMBR:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. MC: Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **TSS:** Conceptualization, Investigation, Methodology, Writing – original draft, Writing – review & editing. **CEF:** Conceptualization, Methodology, Project administration, Writing – review & editing. **LGOB:** Writing – review & editing. **EO:** Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Berman L, Berman J, Miles M, Pollets D, Powell JA. Genital self-image as a component of sexual health: relationship between genital self-image, female sexual function, and quality of life measures. J Sex Marital Ther. 2003;29(Suppl 1):11-21. https://doi.org/10.1080/713847124
2. Yulevitch A, Czamanski-Cohen J, Segal D, Ben-Zion I, Kushnir T. The vagina dialogues: genital self-image and communication with physicians about sexual dysfunction and dissatisfaction among Jewish patients in a women's health clinic in southern Israel. J Sex Med. 2013;10(12):3059-68. https://doi.org/10.1111/jsm.12322
3. Goodman MP, Placik OJ, Matlock DL, Simopoulos AF, Dalton TA, Veale D, et al. Evaluation of body image and sexual satisfaction in women undergoing female genital plastic/cosmetic surgery. Aesthet Surg J. 2016;36(9):1048-57. https://doi.org/10.1093/asj/sjw061
4. Ellibes Kaya A, Yassa M, Dogan O, Basbug A, Pulatoglu C, Caliskan E. The Female Genital Self-Image Scale (FGSIS): cross-cultural adaptation and validation of psychometric properties within a Turkish population. Int Urogynecol J. 2019;30(1):89-99. https://doi.org/10.1007/s00192-018-3688-1
5. Rowen TS, Gaither TW, Shindel AW, Breyer BN. Characteristics of genital dissatisfaction among a nationally representative sample of U.S. women. J Sex Med. 2018;15(5):698-704. https://doi.org/10.1016/j.jsxm.2018.03.004
6. Laan E, Martoredjo DK, Hesselink S, Snijders N, Lunsen RHW. Young women's genital self-image and effects of exposure to pictures of natural vulvas. J Psychosom Obstet Gynaecol. 2017;38(4):249-55. https://doi.org/10.1080/0167482X.2016.1233172
7. Zielinski RE, Kane-Low L, Miller JM, Sampselle C. Validity and reliability of a scale to measure genital body image. J Sex Marital Ther. 2012;38(4):309-24. https://doi.org/10.1080/0092623X.2011.569639. Erratum in: J Sex Marital Ther. 2013;39(4):392.
8. Handelzalts JE, Yaakobi T, Levy S, Peled Y, Wiznitzer A, Krissi H. The impact of genital self-image on sexual function in women with pelvic floor disorders. Eur J Obstet Gynecol Reprod Biol. 2017;211:164-8. https://doi.org/10.1016/j.ejogrb.2017.02.028
9. Digesu GA, Tailor V, Preti M, Vieira-Baptista P, Tarcan T, Stockdale C, et al. The energy based devices for vaginal "rejuvenation," urinary incontinence, vaginal cosmetic procedures, and other vulvo-vaginal disorders: an international multidisciplinary expert panel opinion. Neurourol Urodyn. 2019;38(3):1005-8. https://doi.org/10.1002/nau.23927
10. Barbara G, Facchin F, Buggio L, Alberico D, Frattaruolo MP, Kustermann A. Vaginal rejuvenation: current perspectives. Int J Womens Health. 2017;9:513-9. https://doi.org/10.2147/IJWH.S99700
11. DeMaria AL, Hollub AV, Herbenick D. The Female Genital Self-Image Scale (FGSIS): validation among a sample of female college students. J Sex Med. 2012;9(3):708-18. https://doi.org/10.1111/j.1743-6109.2011.02620.x
12. Mohammed GFAE, Hassan H. Validity and reliability of the arabic version of the female Genital Self-Image Scale. J Sex Med. 2014;11:1193-200. https://doi.org/10.1111/jsm.12494
13. Bump RC, Mattiasson A, Bo K. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. Am J Obstet Gynecol. 1996;175:11-7.
14. Beaton DE, Bombardier C, Guillemin F, Ferraz MB. Guidelines for the process of cross-cultural adaptation of self-report

measures. Spine (Phila Pa 1976). 2000;25(24):3186-91. https://doi.org/10.1097/00007632-200012150-00014

15. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika. 1951;16(3):297-334.
16. Cicchetti D. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess. 1994;6(4):284-90.
17. Danielsen AK, Pommergaard HC, Burcharth J, Angenete E, Rosenberg J. Translation of questionnaires measuring health related quality of life is not standardized: a literature based research study. PLoS One. 2015;10(5):e0127050. https://doi.org/10.1371/journal.pone.0127050
18. El-Azab AS, Abd-Elsayed AA, Imam HM. Patient reported and anatomical outcomes after surgery for pelvic organ prolapse. Neurourol Urodyn. 2009;28:219-24. https://doi.org/10.1002/nau.20626
19. Pasquali L. Psicometria. Rev Esc Enferm USP. 2009;43:992-9.
20. Pakpour AH, Zeidi IM, Ziaeiha M, Burri A. Cross-cultural adaptation of the Female Genital Self-Image Scale (FGSIS) in Iranian female college students. J Sex Res. 2014;51(6):646-53. https://doi.org/10.1080/00224499.2013.821441
21. Zielinski R, Miller J, Low LK, Sampselle C, DeLancey JO. The relationship between pelvic organ prolapse, genital body image, and sexual health. Neurourol Urodyn. 2012;31(7):1145-8. https://doi.org/10.1002/nau.22205
22. Fudge MC, Byers S. "I Have a Nice Gross Vagina": understanding young women's genital self-perceptions. J Sex Res. 2017;54(3):351-61. https://doi.org/10.1080/00224499.2016.1155200
23. Silva Gomes TB, Brasil CA, Barreto APP, Ferreira RS, Berghmans B, Lordelo P. Female genital image: is there a relationship with body image? Turk J Obstet Gynecol. 2019;16(2):84-90. https://doi.org/10.4274/tjod.galenos.2019.49799
24. Sharp G, Tiggemann M. Educating women about normal female genital appearance variation. Body Image. 2016;16:70-8. https://doi.org/10.1016/j.bodyim.2015.11.006
25. Moroni RM, Lara LAS, Ferreira CHJ, Constantino MM, Brito LGO. Assessment of body image, sexual function et al (BIPOP) questionnaire. J Sex Med. 2019;16(1):126-36. https://doi.org/10.1080/00224499.2016.1155200

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20251444

# Psychometric properties of the Brazilian 14-Item Health Literacy Scale for people living with mental disorders: a cross-sectional study

Marcos Renan Barbosa[1], Lucas Machado Maracci[2], Marília Jesus Batista de Brito Mota[3], Eliana Elisa Rehfeld Gheno[1], Márcio Strassburger Junior[1], Mariana Fröhlich Alievi[1], Olga Maria Martins de Sousa Valentim[4,5], Adriane Cristina Bernat Kolankiewicz[1]*

## SUMMARY

**INTRODUCTION:** There is no evidence of validity and reliability for an instrument capable of comprehensively assessing health literacy (including its Functional, Communicative, and Critical dimensions) among Brazilian adults living with mental disorders.

**OBJECTIVE:** The aim of this study was to evaluate the validity, reliability, and evidence of divergent validity of the Brazilian 14-Item Health Literacy Scale in individuals with chronic mental disorders.

**METHODS:** This cross-sectional psychometric study involved patients receiving care at the Psychosocial Care Center in the city of Ijuí, southern Brazil, between April and October 2023. Eligible participants were clinically stable adults with severe and persistent mental disorders; individuals with intellectual disabilities or legal incapacity were excluded. Data collection was conducted by trained students, who read and explained the items of each instrument to the participants. In total, the sample comprised 444 patients. The Brazilian version of the 14-Item Health Literacy Scale, previously translated, cross-culturally adapted, and validated for Brazilian Portuguese, was used to assess health literacy and to perform the psychometric analysis.

**RESULTS:** Good internal consistency was demonstrated, with Cronbach's alpha and McDonald's omega coefficients of 0.83 (95%CI 0.78–0.88) and 0.75 (95%CI 0.68–0.83), respectively. The 14-Item Health Literacy Scale presented acceptable fit indices for the Comparative Fit Index, Tucker-Lewis Index, and Normed Fit Index, with values of 0.938, 0.923, and 0.904, respectively. Divergent validity was verified by the lack of statistically significant correlations and by weak correlations (r<0.300) with the Social Support Scale.

**CONCLUSION:** The 14-Item Health Literacy Scale presented adequate psychometric properties and showed evidence of validity and acceptable reliability for assessing the health literacy of Brazilians living with mental disorder.

**KEYWORDS:** Health literacy. Mental disorders. Validation study. Brazil. Mental health.

## INTRODUCTION

Mental disorders (MDs) are syndromes resulting from psychological, biological, or developmental dysfunctions that impair cognition, emotions, and behavior[1]. The prevalence of common MDs in Brazil is approximately 20%[2], and an estimated 5 million individuals have severe and persistent MDs[3]. Studies conducted in Psychosocial Care Centers (CAPS) indicate a higher frequency of schizophrenic, schizotypal, and delusional disorders[4]. Due to their chronicity, associated disabilities, and socioeconomic impact, MDs represent a significant public health issue[5], affecting individuals, families, and society[6]. Furthermore, due to the chronicity and disability caused by MDs and the consequent impact on the individual and collective economy, these disorders have become an important public health issue to be discussed[7].

Health literacy (HL) refers to the ability to obtain, understand, and use health information[8], in order to make decisions and take actions that promote one's own well-being and that of others, as well as the ability to identify and implement actions aimed at promoting health and improving quality of life. Higher HL levels are associated with better autonomy and treatment adherence[9-11], whereas low HL impairs health management and is associated with lower utilization of health services and greater difficulty in autonomously managing health conditions[12,13].

HL comprises three dimensions: Functional, Communicative, and Critical. The Functional dimension refers to basic reading and writing skills used in everyday situations; the Communicative dimension involves more advanced cognitive and social skills;

[1]Universidade Regional do Noroeste do Estado do Rio Grande do Sul, Graduate Program in Comprehensive Health Care – Ijuí (RS), Brazil.
[2]Universidade Federal de Santa Maria, Graduate Program in Dental Sciences – Santa Maria (RS), Brazil.
[3]Faculdade de Ciências Médicas da Santa Casa de São Paulo, Graduate Program in Dentistry – São Paulo (SP), Brazil.
[4]University of Lisbon, School of Nursing, Nursing Research Innovation and Development Centre of Lisbon (CIDNUR) – Lisbon, Portugal.
[5]RISE-Health, Nursing School of Porto – Porto, Portugal.
*Corresponding author: adri.saudecoletiva@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 29, 2026. Accepted on July 06, 2026.
Scientific Editor: Roseli Nomura

lastly, the Critical dimension refers to a set of cognitive and social skills that enable individuals to critically analyze information and apply it to situations in their daily lives[14]. There are several instruments in the literature that seek to assess Functional HL[15]. However, few instruments are capable of assessing the different dimensions of HL, such as the 14-Item Health Literacy Scale (HLS-14)[16], which is an instrument validated for use in adults regardless of health status. In 2020, it was validated and translated into Brazilian Portuguese after application to a cohort of Brazilian adults, and demonstrated good internal consistency[17].

In Brazil, HL among people with MDs remains understudied[18,19], as there is no validated instrument available to measure this construct in adults living with this condition, which limits both research and the assessment of this important variable in this population. Furthermore, most studies have focused exclusively on the assessment of functional HL[15,20]. Given the high prevalence of MDs[2] and the need for a valid and reliable instrument to assess HL from the perspective of people living with MDs, this study aimed to evaluate the validity, reliability, and evidence of divergent validity of the Brazilian HLS-14 in individuals with chronic mental disorders.

## METHODS

The study was approved by the Research Ethics Committee of the Regional University of Northwest Rio Grande do Sul (protocol no. 5.966.864/2022), in accordance with the 1964 Helsinki Declaration and its subsequent amendments.

This is a cross-sectional psychometric study conducted at a Psychosocial Care Center (CAPS) located in the city of Ijuí, southern Brazil, that has 83,947 inhabitants, integrated into the matrix project "Health Literacy from the Perspective of People with Mental Disorders and Health Professionals: A Mixed-Methods Study." Participants were users with severe and persistent mental disorders in stable clinical condition. Individuals with intellectual disability/mental retardation, judicial interdiction, and/or International Classification of Diseases, 10th Revision (ICD-10) diagnoses F70–F79 were excluded.

Data collection occurred between April and October 2023, conducted by previously trained students. Sample size calculation[21] considered a 90% confidence level and a 9.5% margin of error for a population of 720 users, resulting in a minimum of 432 participants. Participants were recruited through convenience sampling until the target sample size was achieved. Eligibility was assessed according to the inclusion and exclusion criteria. Upon arriving at the facility, patients were invited to participate in the study. After agreeing to participate, they signed the Informed Consent Form and were then escorted to a private room to ensure confidentiality and privacy.

The following instruments were applied: a sociodemographic questionnaire; the HLS-14[17]; and the Medical Outcomes Study Social Support Scale (MOS-SSS)[22]. The HLS-14 is a validated instrument for assessing HL in Brazil[17]; it comprises 14 items (5 assessing Functional HL, 5 assessing Communicative HL, and 4 assessing Critical HL) rated on a five-point Likert scale that indicates how much the participant agrees or disagrees with the item ("strongly disagree" to "strongly agree"). For items 1–5 (Functional dimension), scores range from 1 to 5, with agreement indicating lower HL; in contrast, for items 6–10 (Communicative dimension) and 11–14 (Critical dimension), scoring is reversed (5 to 1), as agreement is associated with higher HL. The MOS-SSS is a multidimensional instrument used to assess users' perceived social support (SS), consisting of 19 items distributed across five dimensions: tangible support, affectionate support, emotional support, informational support, and positive social interaction. Responses are rated on a five-point frequency scale reflecting perceived availability of support: always (5), most of the time (4), sometimes (3), rarely (2), and never (1)[22].

Statistical analyses were performed using Statistical Package for the Social Sciences (SPSS) 25.0 and Analysis of Moment Structures (AMOS), through Confirmatory Factor Analysis (CFA) within Structural Equation Modeling. Outliers (Mahalanobis distance), normality (skewness, kurtosis, and Mardia's coefficient), maximum likelihood estimation, and model fit indices ($\chi^2$, Root Mean Square Residual [RMR], Root Mean Square Error of Approximation [RMSEA], Goodness-of-Fit Index [GFI], Comparative Fit Index [CFI], Normed Fit Index [NFI], and Tucker-Lewis Index [TLI]) were assessed. Reliability was evaluated using Cronbach's alpha and McDonald's omega (0.70–0.95). Convergent/divergent validity evidence was examined through linear associations between HLS-14 and MOS-SSS.

## RESULTS

In total, 512 eligible individuals were invited to participate, 68 (13%) of whom declined, resulting in a final sample of 444 participants (87% response rate). Of these, 75.9% were women, 76.5% were white, and 31.8% were between 55 and 64 years old. Most had a household income of up to two minimum wages (77.4%) and lived with others—including partners, children, or other people (83.3%). Medical records revealed diagnoses of depressive disorders (61%), anxiety disorders (15%), bipolar mood disorder (10%), schizophrenia, schizotypal disorders, delusional disorders (11%), personality and behavioral disorders (1%), and other disorders (1%).

Regarding the validation of the HLS-14, the 14 items were distributed across three dimensions. The validation of the scale structure, performed using CFA, initially considered the basic assumptions for applying the technique and identified evidence of univariate normality, with absolute values of skewness and kurtosis estimated below 3 and 5, respectively. However, the Mardia coefficient (Coefficient=56.473) indicated an asymmetric multivariate distribution. Given this limitation, the structural model was estimated using the Maximum Likelihood method.

According to the results obtained, based on the weight of the individual item loadings, the 14 items of the scale saturated with a magnitude greater than 0.400 ($p<0.001$). The items with the highest loadings are Q8 ($\lambda=0.806$), Q7 ($\lambda=0.762$), and Q13 ($\lambda=0.765$), while items Q10 ($\lambda=0.443$); Q6 ($\lambda=0.539$); and Q11 ($\lambda=0.568$) were items with less weight and less explanatory power, although significant.

Regarding correlations between dimensions, low estimates were detected for the Functional dimension compared to the Communicative (r=-0.340) and Critical (r=-0.250) dimensions. The results with negative coefficients are noteworthy, as the Functional dimension presents reversed items in its assessment. However, between the Communicative and Critical dimensions, the correlation was positive and strong (r=0.710), indicating a degree of dependence between the items. This finding suggests the possibility of these items merging into a single factor, as supported by both the model's quality indices and the dimensions' internal consistency estimates.

Regarding the results achieved by the three-dimensional model structure, quality indicators (or modification indices) were observed, which allow us to identify how well the sample of 444 individuals conformed to the instrument's original structure. Table 1 presents the goodness-of-fit indices for the three-dimensional model defined by CFA on the HLS-14.

An initial assessment of the model's parsimonious fit to the sample ($\chi^2/df$) indicated minimal discrepancy between the observed and estimated covariance matrices, with a goodness-of-fit value below 4.00 ($[\chi^2/df]=2.626<4.0$). Furthermore, regarding the sample's fit to the scale's theoretical structure, we investigated whether the degree of disagreement between the model and the sample covariance was within the acceptance limits of an adequate fit, using the RMSEA. Also, according to the result obtained (RMSEA=0.061<0.08), the observed disagreements did not compromise the credibility of the model. Comparative indices (CFI=0.938; TLI=0.923; NFI=0.904) indicated satisfactory fit.

As shown in Table 1, the absolute fit indices GFI (Goodness-of-Fit Index) and AGFI (Adjusted Goodness-of-Fit Index) presented values above the minimum acceptable threshold (0.900).

Regarding the relative indexes, which assess how much better the proposed model fits the data compared to the base model generated by AMOS (hypothetical model), the estimates for TLI and CFI were above the acceptable cutoff point, with estimates of 0.923 and 0.938, respectively. The NFI, which investigates the proportion of the difference between the Chi-square value of the proposed model in relation to a null model, was estimated at 0.904 (>0.900), indicating that, for the sample studied, the model presents representative correlations that enable the proposed dimensions.

As for reliability, the Omega ($\omega$) coefficient for the total HLS-14 was 0.755 (95%CI 0.678–0.831), indicating satisfactory reliability ($\omega>0.700$), as shown in Table 2. Regarding the Cronbach's alpha estimate, this coefficient reached 0.831 (95%CI 0.781–0.880). It was observed that the participants presented low Functional literacy (2.83) and high Critical literacy (3.62) and Communicative literacy (3.57).

Construct equivalence involves demonstrating divergent validity through the relationship of the main scale with other measures. In this study, the HLS-14 instrument was compared with the MOS-SSS questionnaire. As observed in Table 3, divergent validity was evidenced by the absence of significant correlations, as well as the presence of significant correlations of weak magnitude (r<0.300) between the compared dimensions.

**Table 1.** Goodness-of-fit indices obtained from the confirmatory factor analysis of the original three-factor structure of the Brazilian 14-Item Health Literacy Scale in a sample of adults with mental disorders (n=444).

| Data adequacy indices | Confirmatory factor model of the HLS-14 |
|---|---|
| | Estimators |
| Normed $\chi^2$ ($\chi^2/df$) | **2.626** |
| RMSEA [95%CI] | **0.061 [0.050–0.071]** |
| Fit indices: comparative indices | |
| AGFI | **0.917** |
| GFI | **0.941** |
| CFI | **0.938** |
| TLI | **0.923** |
| NFI | **0.904** |
| Predictive fit indices | |
| AIC | **2.006** |
| RMR (SRMR) | **0.056 (0.043)** |

AGFI: Adjusted Goodness-of-Fit Index; AIC: Akaike Information Criterion; CFI: Comparative Fit Index; GFI: Goodness-of-Fit Index; df: degrees of freedom; CI: confidence interval; HLS-14: 14-Item Health Literacy Scale; NFI: Normed Fit Index; RMR: Root Mean Square Residual; RMSEA: Root Mean Square Error of Approximation; SRMR: Standardized Root Mean Square Residual; TLI: Tucker-Lewis Index; $\chi$2: chi-square; $\chi$2/df: normed chi-square. The bold values indicate the estimated goodness-of-fit indices obtained from the confirmatory factor analysis.

## DISCUSSION

This cross-sectional study demonstrated that the HLS-14 instrument has valid and reliable psychometric properties for measuring HL in people living with MDs. This is the first Brazilian study employing the HLS-14 in this population. Participants presented low Functional literacy and high Critical and Communicative literacy.

Validating a tool that allows the assessment of HL for this specific population is important, given the considerable prevalence of MDs in Brazil[2,4] and worldwide[23], as well as the important relationship between HL and autonomy and treatment adherence[10,11]. Furthermore, most studies on HL of people living with MDs are limited to assessing the Functional dimension, providing only a partial understanding, since HL is currently a multidimensional construct[15]. Thus, the validation of the HLS-14 is important, as it may play a valuable role in identifying impaired dimensions and planning targeted interventions by assessing all three dimensions.

Confirmatory factor analysis demonstrated adequate model fit, consistent with previous validations[16,17]. The RMSEA was 0.061, indicating an acceptable model fit[24]. Based on the goodness-of-fit indicators obtained, it is noteworthy that the evaluation of model fit generally requires that at least three fit indices meet or exceed the recommended thresholds for adequate fit. In this regard, the three-factor model showed a substantial number of satisfactory fit indicators, providing evidence of the high quality of the factor structure tested for the HLS-14.

Reliability was adequate, with an overall Cronbach's alpha of 0.83 (>0.70), consistent with previous studies (0.81–0.82)[16,17]. McDonald's omega was 0.755, confirming internal consistency. Dimension-specific alphas were 0.83 (Functional), 0.90 (Communicative), and 0.77 (Critical), comparable to the literature[16,17], with slight variation in the Functional domain compared to the Brazilian validation study[17].

The lowest mean was observed for Functional HL (2.83), whereas Communicative (3.57) and Critical (3.62) dimensions showed similar values. Greater variability was found in the Functional domain. These values suggest that, while the study participants are reasonably good at understanding and using health information, they may face difficulties in more basic aspects, such as reading and initial comprehension of the information. Furthermore, considering that the functional dimension seeks to assess basic reading and writing skills in everyday situations, the low average score in this dimension may be explained by the

**Table 2.** Measures of central tendency, variability, and internal consistency (McDonald's omega and Cronbach's alpha) for the dimensions of the Brazilian 14-Item Health Literacy Scale in adults with mental disorders (n=444).

| Dimensions | HLS-14 estimates | | | | | | |
|---|---|---|---|---|---|---|---|
| | Mean | Standard deviation | Quartiles | | | Reliability (95%CI) | |
| | | | Q1 | Median Q2 | Q3 | ωʃ | αC |
| Functional HL | 2.83 | 0.86 | 2.20 | 2.80 | 3.40 | 0.814 (0.740–0.887) | 0.826 (0.784–0.867) |
| Communicative HL | 3.57 | 0.61 | 3.20 | 3.60 | 4.00 | 0.892 (0.838–0.945) | 0.899 (0.845–0.952) |
| Critical HL | 3.62 | 0.66 | 3.25 | 3.75 | 4.00 | 0.720 (0.714–0.725) | 0.767 (0.700–0.833) |
| Total | 3.32 | 0.53 | 3.00 | 3.36 | 3.57 | 0.755 (0.678–0.831) | 0.831 (0.781–0.880) |

HLS-14: 14-Item Health Literacy Scale; CI: confidence interval; HL: health literacy; αC: Cronbach's alpha; ωʃ: McDonald's Omega (internal consistency with BCa bootstrap confidence intervals, 1,000 samples).

**Table 3.** Correlation coefficients between perceived social support (Medical Outcomes Study Social Support Scale) dimensions and health literacy (14-Item Health Literacy Scale) dimensions in adults with mental disorders (n=444).

| MOS-SSS | HLS-14 | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Functional literacy | | Communicative literacy | | Critical literacy | | Total | |
| | r | p | R | p | r | p | r | p-value |
| Material support | -0.008 | 0.872 | 0.046 | 0.329 | -0.074 | 0.121 | -0.011 | 0.810 |
| Affective support | -0.009 | 0.843 | 0.028 | 0.553 | -0.055 | 0.247 | -0.013 | 0.782 |
| Emotional support | 0.059 | 0.213 | 0.222**q | 0.000 | 0.104* | 0.028 | 0.163** | 0.001 |
| Informational support | 0.002 | 0.959 | 0.162** | 0.001 | 0.046 | 0.334 | 0.085 | 0.073 |
| Emotional/informational support | 0.033 | 0.492 | 0.203** | 0.000 | 0.079 | 0.095 | 0.131** | 0.006 |
| Positive social interaction | 0.030 | 0.533 | 0.186** | 0.000 | 0.141** | 0.003 | 0.145** | 0.002 |

HLS-14: 14-Item Health Literacy Scale; MOS-SSS: Medical Outcomes Study Social Support Scale; r: factor correlation coefficient. *Statistically significant correlation at p<0.05. **Statistically significant correlation at p<0.01.

sample in this study being composed mostly of people with low education level (between incomplete and complete primary education) and low socioeconomic status, factors associated with poorer HL[12]. These findings also corroborate the results of Gheno et al., which also found an influence of skin color and age on HL[21].

All factor loadings were >0.400, demonstrating the explanatory strength of the items and supporting their retention in the instrument. However, unlike previous studies, a borderline factor loading was observed for item 10 ("I apply the information I receive, through self-assessment, to my lifestyle and daily routine."), although it remained acceptable and relevant. This lower factor loading, compared with the other items, may be explained by the greater variability in responses, which could be related to the complexity of the statement and the resulting difficulty in its comprehension. Furthermore, cultural particularities and the clinical context of the participants should be taken into account, given that the study population comprised individuals with different mental disorder diagnoses.

Regarding the correlation among the three dimensions of HL, it was verified a high correlation between Communicative and Critical dimensions, and a low correlation between these and the Functional dimension, as previously reported[16,17], reinforcing the multidimensional nature of HL. This finding suggests that basic information collected from the Functional dimension cannot replace data from the Communicative and Critical dimensions, which assess more advanced knowledge. Therefore, it reinforces the importance of using an instrument that not only assesses the Functional dimension but also takes into account the multidimensionality of literacy.

Divergent validity was confirmed by the absence of significant correlations or weak correlations (r<0.300) with the MOS-SSS[22]. Furthermore, comparison of this finding with previous studies was not possible, as no studies assessing divergent validity between these instruments were identified in the literature.

To our knowledge, this is the first study to validate the HLS-14 for use among people living with MDs. Strengths include the use of a previously validated instrument[17], brief administration, and a large sample. However, some limitations should be considered. The questionnaire was administered by examiners; therefore, the HLS-14 items were subject to explanations and clarifications provided to participants, which may have influenced their understanding of the items and, consequently, their responses. The sample consisted of individuals receiving care at a health service, characterizing a convenience sample that was not selected using statistical sampling procedures. In addition, it included patients with different MD diagnoses, which may have influenced the psychometric performance of the HLS-14. Lastly, the cross-sectional design did not allow the assessment of test-retest reliability, preventing the evaluation of the tool's reproducibility over time. Therefore, future studies using random and representative population-based samples, as well as longitudinal designs, are recommended to confirm the findings of the present study.

The findings of the present study show that the HLS-14 maintained adequate psychometric properties in individuals with MDs, which is of extreme importance since this population often faces challenges in accessing, understanding, and using health information[25]. Reliable assessment of HL may help find needs support care that focuses on the patient and direct efforts to improve treatment adherence, self-care and overall health[15]. Thus, the HLS-14 can play an important role in identifying needs and improving health outcomes for individuals living with MDs, offering an assessment not only of Functional HL, but of the Communicative and Critical dimensions as well.

## CONCLUSION

The HLS-14 showed evidence of validity and acceptable reliability for assessing HL in Brazilian adults and older adults with severe and persistent mental disorders. The instrument demonstrated acceptable internal consistency, adequate fit to the proposed three-factor structure, and evidence of divergent validity, supporting its potential use in this population.

## AUTHORS' CONTRIBUTIONS

**MRB:** Conceptualization, Investigation, Methodology, Resources, Visualization, Writing – original draft, Writing – review & editing. **LMM:** Methodology, Visualization, Writing – original draft, Writing – review & editing. **MJBBM:** Methodology, Visualization, Writing – original draft, Writing – review & editing. **EERG:** Conceptualization, Investigation, Methodology, Resources, Visualization, Writing – original draft, Writing – review & editing. **MSJ:** Methodology, Visualization, Writing – original draft, Writing – review & editing. **MFA:** Conceptualization, Investigation, Methodology, Resources, Visualization, Writing – original draft, Writing – review & editing. **OMMSV:** Methodology, Visualization, Writing – original draft, Writing – review & editing. **ACBK:** Conceptualization, Investigation, Methodology, Resources, Visualization, Writing – original draft, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. World Health Organization. International Classification of Diseases 11th Revision (ICD-11). World Health Organization; 2020. Available from: https://icd.who.int
2. Brunoni AR, Suen PJC, Bacchi PS, Razza LB, Klein I, Santos LA, et al. Prevalence and risk factors of psychiatric symptoms and diagnoses before and during the COVID-19 pandemic: findings from the ELSA-Brasil COVID-19 mental health cohort. Psychol Med. 2023;53(2):446-57. https://doi.org/10.1017/S0033291721001719
3. Lopes CS. How is Brazilian's mental health? The importance of birth cohorts for better understanding the problem. Cad Saude Publica. 2020;36(2):e00005020. https://doi.org/10.1590/0102-311X00005020
4. Barbosa CG, Meira PRM, Nery JS, Gondim BB. Epidemiological profile of the users of a Psychosocial Care Center. SMAD, Rev Eletrônica Saúde Mental Álcool Drog. 2020;16(1):1-8. https://doi.org/10.11606/issn.1806-6976.smad.2020.156687
5. Karambelas GJ, Filia K, Byrne LK, Allott KA, Jayasinghe A, Cotton SM. A systematic review comparing caregiver burden and psychological functioning in caregivers of individuals with schizophrenia spectrum disorders and bipolar disorders. BMC Psychiatry. 2022;22:422. https://doi.org/10.1186/s12888-022-04069-w
6. Alemu WG, Due C, Muir-Cochrane E, Mwanri L, Azale T, Ziersch A. Quality of life among people living with mental illness and predictors in Africa: a systematic review and meta-analysis. Qual Life Res. 2024;33:1191-209. https://doi.org/10.1007/s11136-023-03525-8
7. Arias D, Saxena S, Verguet S. Quantifying the global burden of mental disorders and their economic value. EClinicalMedicine. 2022;54:101675. https://doi.org/10.1016/j.eclinm.2022.101675
8. Sousa RA. Literacia para a saúde: habilidades para lidar com as informações sobre saúde podem ajudar a construir novos caminhos na saúde pública. RECIIS. 2022;16:175-87. https://doi.org/10.29397/reciis.v16i1.3221
9. Lor M, Koleck TA, Bakken S, Yoon S, Dunn Navarra AM. Association between health literacy and medication adherence among hispanics with hypertension. J Racial Ethn Health Disparities. 2019;6:517-24. https://doi.org/10.1007/s40615-018-00550-z
10. Rossato RO, Rocha SHDN. A importância da adesão ao tratamento e o letramento em saúde: uma revisão de literatura. Braz J Health Rev. 2020;3:19672-8. https://doi.org/10.34119/bjhrv3n6-343
11. Suhail M, Saeed H, Saleem Z, Younas S, Hashmi FK, Rasool F, et al. Association of health literacy and medication adherence with health-related quality of life (HRQoL) in patients with ischemic heart disease. Health Qual Life Outcomes. 2021;19:118. https://doi.org/10.1186/s12955-021-01761-5
12. Huang CH, Talley PC, Lin CW, Huang RY, Liu IT, Chiang IH, et al. Factors associated with low health literacy among community-dwelling women in Taiwan. Women Health. 2020;60:487-501. https://doi.org/10.1080/03630242.2019.1662872
13. Shahid R, Shoker M, Chu LM, Frehlick R, Ward H, Pahwa P. Impact of low health literacy on patients' health outcomes: a multicenter cohort study. BMC Health Serv Res. 2022;22(1):1148. https://doi.org/10.1186/s12913-022-08527-9
14. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. Health Promot Int. 2000;15:259-67. https://doi.org/10.1093/heapro/15.3.259
15. Degan TJ, Kelly PJ, Robinson LD, Deane FP, Smith AM. Health literacy of people living with mental illness or substance use disorders: a systematic review. Early Interv Psychiatry. 2021;15(6):1454-69. https://doi.org/10.1111/eip.13090
16. Suka M, Odajima T, Kasai M, Igarashi A, Ishikawa H, Kusama M, et al. The 14-item health literacy scale for Japanese adults (HLS-14). Environ Health Prev Med. 2013;18:407-15. https://doi.org/10.1007/s12199-013-0340-z
17. Batista MJ, Marques ACP, Silva Junior MF, Alencar GP, Sousa MDLR. Translation, cross-cultural adaptation and psychometric evaluation of Brazilian Portuguese version of the 14-item Health Literacy Scale. Cien Saude Colet. 2020;25(7):2847-57. https://doi.org/10.1590/1413-81232020257.22282018
18. Peres PCN, Pessoa KR, Bernuci MP, Massuda EM, Yamaguchi MU. Literacia em saúde no brasil: estudo cienciométrico. Encicl Biosfera. 2017;14:1589. https://doi.org/10.18677/EnciBio_2017A132
19. Moreira WC, Sousa AR, Cardoso RDSS, Queiroz AM, Oliveira MAF, Sequeira CADC. COVID-19 in Brazil: are there any differences in Mental Health Literacy between young and aged men? Rev Lat Am Enfermagem. 2022;30:e3603. https://doi.org/10.1590/1518-8345.5651.3603
20. Pavão ALB, Werneck GL. Health literacy in low- and middle-income countries: a systematic review. Cien Saude Colet. 2021;26(9):4101-14. https://doi.org/10.1590/1413-81232021269.05782020
21. Gheno EER, Rodrigues FO, Kunkel GK, Nogaro A, Anastácio ZFC, Kolankiewicz ACB. Association between sociodemographic variables and health literacy levels among individuals with mental disorders in Brazil. J Health Literacy. 2024;9(3):31-44.
22. Griep RH, Chor D, Faerstein E, Werneck GL, Lopes CS. Construct validity of the Medical Outcomes Study's social support scale adapted to Portuguese in the Pró-Saúde Study. Cad Saude Publica. 2005;21(3):703-14. https://doi.org/10.1590/s0102-311x2005000300004
23. Wu Y, Wang L, Tao M, Cao H, Yuan H, Ye M, et al. Changing trends in the global burden of mental disorders from 1990 to 2019 and predicted levels in 25 years. Epidemiol Psychiatr Sci. 2023;32:e63. https://doi.org/10.1017/S2045796023000756
24. Kline RB. Principles and practice of structural equation modeling. 5th ed. New York (NY): Guilford Press; 2023.
25. Mantell PK, Baumeister A, Christ H, Ruhrmann S, Woopen C. Peculiarities of health literacy in people with mental disorders: a cross-sectional study. Int J Soc Psychiatry. 2020;66(1):10-22. https://doi.org/10.1177/0020764019873683

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20252134

# Sarcopenia, its components and fracture risk in people living with human immunodeficiency virus

Edivaldo Cangussu Meira Junior[1], Jaine Alves Ximenez[2], Marcia Midori Shinzato[1], Maria Claudia Bernardes Spexoto[2]*

## SUMMARY

**OBJECTIVE:** The aim of this study was to evaluate the relationship between sarcopenia and its components with fracture risk in people living with human immunodeficiency virus, and to assess the association with sociodemographic and clinical characteristics.

**METHODS:** This cross-sectional study was conducted with people living with human immunodeficiency virus undergoing antiretroviral therapy in the state of Mato Grosso do Sul, Brazil. Sarcopenia was defined according to the European Working Group on Sarcopenia in Older People criteria, assessing muscle strength (handgrip strength), skeletal muscle mass (skeletal muscle mass index), and physical performance (gait speed). The 10-year risk of major osteoporotic and hip fracture was estimated using the Brazilian version of the Fracture Risk Assessment Tool (FRAX®). Statistical analyses included chi-square tests, analysis of variance, and Spearman correlation ($p<0.05$).

**RESULTS:** Among 105 participants (44.6±13.5 years), the prevalence of confirmed or severe sarcopenia was 6.7%. No association was found between sarcopenia and the risk of major fracture ($p=0.471$) or hip fracture ($p=0.254$). However, handgrip strength was correlated with the risk of major fracture ($\rho=-0.21$; $p=0.029$) and hip fracture ($\rho=-0.25$; $p=0.010$). SMII showed moderate correlations with major fracture risk ($\rho=-0.44$; $p<0.01$) and hip fracture risk ($\rho=-0.52$; $p<0.01$). Gait speed was correlated only with hip fracture risk ($\rho=-0.22$; $p=0.025$). Older age ($p<0.001$) and lower body mass index ($p=0.011$) were significant risk factors.

**CONCLUSION:** Sarcopenia as a categorical diagnosis was not associated with fracture risk, whereas lower handgrip strength and skeletal muscle mass index were associated with higher FRAX® estimates, and gait speed only with hip fracture risk. These findings support assessing sarcopenia-related parameters in people living with human immunodeficiency virus.

**KEYWORDS:** HIV. Sarcopenia. Bone fractures. Muscle strength.

## INTRODUCTION

In the global scenario of 2023, 1.3 million new human immunodeficiency virus (HIV) infections were recorded. In Brazil, the detection rate was 17.1 cases per 100,000 inhabitants, contrasting with 22.4 in the state of Mato Grosso do Sul, and even higher rates in the municipality where this study was conducted[1]. Although notification is effective, the increased longevity achieved imposes challenges related to musculoskeletal health.

Sarcopenia, defined by the simultaneous presence of low muscle strength and low muscle mass[2], is prevalent among people living with HIV (PLHIV), a population that experiences accelerated aging[3] often accompanied by frailty[4]. Persistent chronic inflammation[5], combined with the adverse effects of antiretroviral therapy (ART) on musculoskeletal health[6], induces changes in body composition and reductions in bone mineral density (BMD)[7]. Specifically, the use of tenofovir disoproxil fumarate (TDF) is associated with impaired bone metabolism and the exacerbation of sarcopenia[7].

The cumulative effects of these factors, combined with traditional risk factors, significantly increase the risk of fractures in PLHIV, especially among older individuals[8]. In this context, the use of tools such as FRAX® (Fracture Risk Assessment Tool) becomes essential for risk stratification[9]. The present study aimed to evaluate the relationship between sarcopenia and its components (muscle strength, skeletal muscle mass, and physical performance) with fracture risk in PLHIV, as well as to assess the association with sociodemographic and clinical characteristics of this population.

[1]Universidade Federal da Grande Dourados, Faculty of Health Sciences – Dourados (MS), Brazil.
[2]Universidade Federal da Grande Dourados, Faculty of Health Sciences, Graduate Program in Food, Nutrition and Health – Dourados (MS), Brazil.
*Corresponding author: mariaspexoto@ufgd.edu.br
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 08, 2026. Accepted on July 06, 2026.
Scientific Editor: Roseli Nomura

# METHODS

## Design, population, and data collection

This cross-sectional observational study (Strengthening the Reporting of Observational Studies in Epidemiology guidelines—Strengthening the Reporting of Observational Studies in Epidemiology) was conducted in Dourados, MS, Brazil (April 2023 to December 2024) at the University Hospital of the Federal University of Grande Dourados/Hospital Universitário da Universidade Federal da Grande Dourados (HU-UFGD) and at the Specialized Care Service and Testing and Counseling Center (SAE/CTA). Participants were recruited consecutively among eligible PLHIV receiving care in outpatient and inpatient settings during the study period. Eligible individuals were those on ART who were able to undergo the study assessments and provide informed consent. Individuals with complete mobility restrictions, cognitive impairment preventing consent, edema, or an HIV diagnosis established during the hospitalization in which study recruitment occurred were excluded.

Data collection was carried out through the administration of structured forms and the review of medical records (both physical and electronic). Sociodemographic and clinical variables were assessed. The sociodemographic variables included age (years), sex, race/skin color, marital status, and educational level. The clinical variables comprised disease status (asymptomatic, symptomatic, or Aids), history of opportunistic and chronic diseases (yes/no), body mass index (BMI) in kg/m$^2$, CD4+ count (>200 or ≤200 cells/mm$^3$), and ART regimens.

## Assessment of sarcopenia and its components

Sarcopenia was defined according to the revised criteria of the European Working Group on Sarcopenia in Older People (EWGSOP2)[2]. Muscle strength was assessed using handgrip strength (HGS), measured with a hydraulic dynamometer (SAEHAN®, model SH5001), considering the highest value among three trials (cutoff points: <27 kg for men; <16 kg for women). Skeletal muscle mass (SMM) was estimated using the Lee equation[10] and divided by height squared to calculate the skeletal muscle mass index (SMMI), adopting cutoff points established for the Brazilian population (men <8.9 kg/m$^2$; women <6.4 kg/m$^2$)[11]. Physical performance was evaluated using gait speed (GS), with a cutoff value of ≤0.8 m/s$^2$. Participants were categorized as non-sarcopenic, probable sarcopenia, confirmed sarcopenia, or severe sarcopenia, according to the EWGSOP2 algorithm[2]. In this study, confirmed sarcopenia and severe sarcopenia were grouped together and classified as "sarcopenia."

## Fracture risk

The 10-year risks of major osteoporotic fracture (MOF) and hip fracture (HF) were estimated using the FRAX® Brazil tool and expressed as percentages[12]. FRAX® estimates fracture probability based on clinical risk factors, with or without BMD. For descriptive comparisons, FRAX® estimates were categorized according to their distribution in the sample into three quantile-based groups, labeled Q1, Q2, and Q3. For MOF, the categories were Q1 (≤0.40%), Q2 (>0.40% to ≤1.00%), and Q3 (>1.00%). For HF, the categories were Q1 (0.00%), Q2 (>0.00% to ≤0.20%), and Q3 (>0.20%).

## Statistical analysis

The analyses were performed using Statistical Package for the Social Sciences version 22. Normality was assessed using the Kolmogorov-Smirnov test. Descriptive statistics were applied, and Pearson's chi-square or Fisher's exact tests were used to examine associations, while analysis of variance (ANOVA) was employed for multiple comparisons. The correlation between sarcopenia components and FRAX® scores was evaluated using Spearman's coefficient (ρ). A significance level of $p<0.05$ was adopted.

## Ethical considerations

The study was approved by the Research Ethics Committee of UFGD under approval numbers 5.919.928; 6.559.968; and 7.684.435. All participants signed the informed consent form.

# RESULTS

A total of 105 participants (81 outpatients and 24 inpatients) were evaluated after the exclusion of 12 individuals based on clinical criteria and 14 refusals. The sample had a mean age of 44.6±13.6 years, with a slight predominance of males (51.4%). Sociodemographic characteristics and their association with fracture risk are detailed in Table 1. Age was significantly associated with both MOF and HF FRAX® estimates ($p<0.001$ for both), with participants in Q3 being older.

Regarding clinical characteristics (Table 2), the mean BMI was 24.19±6.20 kg/m$^2$ and was associated only with HF FRAX® estimates ($p=0.011$), which participants in Q3 showed lower BMI values.

TDF was used by 75.2% of the participants. The predominant ART regimen was a triple therapy consisting of two nucleos(t)ide reverse transcriptase inhibitors and one integrase inhibitor (66.7%).

The distribution of participants according to sarcopenia status was 78.1% non-sarcopenic, 15.2% probable sarcopenia,

**Table 1.** Sociodemographic variables and their relationship with fracture risk (Dourados, MS, Brazil, 2023–2024).

| Variables | n* | Major osteoporotic fracture risk, FRAX® (%) | | | | | Hip fracture risk, FRAX® (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | %† | Q1‡ | Q2§ | Q3‖ | p¶ value | Q1‡ | Q2§ | Q3‖ | p¶ value |
| Age (years) (Mean±SD**) | 44.6±13.6 | | 34.8±7.5 | 41.8±12.2 | 51.1±13.2 | <0.001 | 35.2±7.8 | 44.9±11.5 | 51.1±14.6 | <0.001 |
| Sex | | | | | | | | | | |
| Male | 54 | 51.4 | 14 | 16 | 24 | 0.439 | 15 | 20 | 19 | 0.410 |
| Female | 51 | 48.6 | 13 | 10 | 28 | | 15 | 13 | 23 | |
| Self-reported race/skin color | | | | | | | | | | |
| Non-white | 75 | 71.4 | 19 | 21 | 35 | 0.459 | 22 | 24 | 29 | 0.906 |
| White | 30 | 28.6 | 8 | 5 | 17 | | 8 | 9 | 13 | |
| Marital status | | | | | | | | | | |
| Married | 34 | 32.4 | 8 | 11 | 15 | 0.122 | 10 | 13 | 11 | 0.201 |
| Single | 46 | 43.8 | 14 | 9 | 23 | | 15 | 12 | 19 | |
| Widowed | 7 | 6.7 | - | - | 7 | | - | 1 | 6 | |
| Separated/divorced | 18 | 17.1 | 5 | 6 | 7 | | 5 | 7 | 6 | |
| Education level | | | | | | | | | | |
| <4 years | 11 | 10.5 | 2 | 3 | 6 | 0.050 | 2 | 4 | 5 | 0.523 |
| 4 to 8 years | 44 | 41.9 | 12 | 4 | 28 | | 12 | 10 | 22 | |
| 8 to 11 years | 29 | 27.6 | 8 | 12 | 9 | | 10 | 11 | 8 | |
| >11 years | 21 | 20.0 | 5 | 7 | 9 | | 6 | 8 | 7 | |

*n: sample size; †%: percentage; ‡Q1, §Q2, and ‖Q3: quantile-based FRAX® risk groups according to the sample distribution, expressed as 10-year fracture probabilities (%); ¶p: Pearson's chi-square test was used for categorical variables, and the ANOVA (analysis of variance) test was used for the numerical variable; **SD: standard deviation; FRAX®: Fracture Risk Assessment Tool. Bold values indicate statistical significance (p<0.05).

**Table 2.** Clinical variables and their relationship with fracture risk (Dourados, MS, Brazil, 2023–2024).

| Variables | n* | %† | Major osteoporotic fracture risk, FRAX® (%) | | | p¶ value | Hip fracture risk, FRAX® (%) | | | p¶ value |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | Q1‡ | Q2§ | Q3‖ | | Q1‡ | Q2§ | Q3‖ | |
| Disease status | | | | | | | | | | |
| Asymptomatic | 32 | 30.5 | 12 | 6 | 14 | 0.384 | 13 | 8 | 11 | 0.332 |
| Symptomatic | 37 | 35.2 | 7 | 9 | 21 | | 8 | 11 | 18 | |
| Aids | 36 | 34.3 | 8 | 11 | 17 | | 9 | 14 | 13 | |
| Opportunistic diseases | | | | | | | | | | |
| Absent | 69 | 65.7 | 19 | 19 | 31 | 0.418 | 21 | 22 | 26 | 0.768 |
| Present | 36 | 34.3 | 8 | 7 | 21 | | 9 | 11 | 16 | |
| Previous chronic diseases | | | | | | | | | | |
| Absent | 70 | 66.7 | 17 | 18 | 35 | 0.881 | 19 | 22 | 29 | 0.879 |
| Present | 35 | 33.3 | 10 | 8 | 17 | | 11 | 11 | 13 | |
| Body mass index¶ (kg/m²) (Mean±SD**) | 24.19±6.20 | | 25.6±5.3 | 24.7±7.8 | 23.2±5.7 | 0.257 | 26.2±5.6 | 25.1±7.0 | 22.1±5.3 | 0.011 |
| CD4+††† (cells/mm³)‡‡‡ (Med§§§) | 342.5 | | | | | 0.588 | | | | 0.491 |
| ≤200 cells/mm³ | 36 | 34.3 | 8 | 11 | 17 | | 9 | 14 | 13 | |
| >200 cells/mm³ | 69 | 65.7 | 19 | 15 | 35 | | 21 | 19 | 29 | |

*n: sample size; †%: percentage; ‡Q1, §Q2, and ‖Q3: quantile-based FRAX® risk groups according to the sample distribution, expressed as 10-year fracture probabilities (%); ¶p: Pearson's chi-square test was used for categorical variables, and the ANOVA (analysis of variance) test was used for the numerical variable; **SD: standard deviation; FRAX®: Fracture Risk Assessment Tool; †††CD4+: CD4+ T-helper lymphocyte count; ‡‡‡2 missing cases; §§§Med: median. Bold values indicate statistical significance (p<0.05).

and 6.7% confirmed or severe sarcopenia. Sarcopenia as a categorical diagnosis was not associated with MOF (p=0.471) or HF risk estimates (p=0.254). In contrast, when sarcopenia-related parameters were analyzed individually, SMMI was inversely correlated with MOF (ρ=-0.44; p<0.01) and HF (ρ=-0.52; p<0.01), and HGS was also inversely correlated with MOF (ρ=-0.21; p=0.029) and HF (ρ=-0.25; p=0.010). GS was inversely correlated only with HF risk (ρ=-0.22; p=0.025) (Table 3).

## DISCUSSION

In this study, sarcopenia as a categorical diagnosis was not associated with MOF or HF risk FRAX® estimates. In contrast, lower SMMI and HGS were correlated with higher MOF and HF estimates, whereas lower GS was correlated only with HF. This pattern may reflect the low prevalence of confirmed/severe sarcopenia and the loss of information when continuous measures are categorized into a composite diagnosis. Older age was associated with both outcomes, and lower BMI with HF.

The lack of an association with sarcopenia in our study contrasts with evidence from the literature. In one cohort, individuals with sarcopenia showed a significantly higher risk of osteoporotic fracture (HR 2.11 [95%CI 1.47–3.04]), with this risk being even higher among women (HR 2.40 [95%CI 1.51–3.81])[13]. Supporting this scenario, recent data indicate that osteosarcopenia—defined as the coexistence of sarcopenia and osteopenia—(with a global prevalence of 18.5%) represents a strong risk factor for fractures (HR 2.13), falls, and mortality[14]. This divergence in our findings may reflect the sample size or specific clinical characteristics of the population studied, which presented advanced immunosuppression (median CD4+ count of 342.5 cells/mm$^3$). The literature indicates that HIV infection and chronic inflammation dysregulate the RANKL (receptor activator of nuclear factor κB ligand)/OPG (osteoprotegerin) pathway, enhancing osteoclastic activity and accelerating bone loss[15]. In addition, the predominant use of TDF in the sample is a well-known factor associated with reduced BMD[5].

Low BMI was significantly associated with HF risk, supporting evidence that low body weight is an independent risk factor, reflecting reduced muscle and bone reserves[16]. Although the use of integrase strand transfer inhibitors (present in the ART regimen of most PLHIV in this sample) has been associated with weight gain in a systematic review[17], maintaining an adequate BMI remains a critical protective factor.

The analysis of sarcopenia components revealed that muscle strength and muscle mass are relevant indicators. Low HGS emerged as a risk predictor, aligning with recent studies that position it as an effective tool for screening bone fragility, independent of BMD[18]. Similarly, GS was correlated with HF risk; evidence suggests that slow GS may increase fracture risk by up to 1.8-fold[16], whereas preserved mobility serves as a protective factor against sarcopenia in PLHIV[19].

Early aging, exacerbated by the interaction between HIV, ART, and comorbidities, accelerates musculoskeletal deterioration[20], which justifies the association observed between older age and fractures.

The main limitations of this study are its cross-sectional design, modest sample size, and inclusion of participants from both outpatient and inpatient settings. Although this approach reflects real-world care, these settings may differ in disease severity, mobility, nutritional status, and muscle function. Given the small number of inpatients, stratified analyses by care setting were not performed. Accordingly, the findings should be interpreted as applying to the overall study population rather than to each care setting separately.

As strengths, the study demonstrated the relevance of sarcopenia components in functional assessment for fracture risk. Moreover, the use of validated tools (EWGSOP2, FRAX®) in an understudied population supports the implementation of multidisciplinary screening routines in PLHIV.

**Table 3.** Correlation between sarcopenia components and fracture risk scores among participants (Dourados, MS, Brazil, 2023–2024).

| Components of sarcopenia | | Major fracture risk | Hip fracture risk |
|---|---|---|---|
| HGS* | ρ† | -0.21 | -0.25 |
| | p‡ | <0.029 | <0.010 |
| SMII§ | ρ† | -0.44 | -0.52 |
| | p‡ | <0.010 | <0.010 |
| GS‖ | ρ† | -0.16 | -0.22 |
| | p‡ | 0.095 | <0.025 |

*HGS: maximum handgrip strength; †ρ: Spearman correlation coefficient; ‡p: significance level; §SMII: skeletal muscle mass index; ‖GS: gait speed.

## CONCLUSION

Although categorical sarcopenia was not associated with FRAX® estimates, lower HGS and SMMI were correlated with higher MOF and HF estimates, while lower GS and BMI were correlated only with HF. Older age was associated with both outcomes. These findings support the assessment of sarcopenia-related parameters in PLHIV.

## ETHICAL APPROVAL

The study was approved by the Research Ethics Committee of the Federal University of Grande Dourados (UFGD) on March 1, 2023, under approval numbers 5.919.928, 6.559.968, and 7.684.435, and all participants signed the informed consent form.

## ACKNOWLEDGMENTS

We thank the National Council for Scientific and Technological Development (CNPq) and the Brazilian Hospital Services Company (Ebserh) for granting the scientific research fellowship, and HU-UFGD for supporting the study.

## AUTHORS' CONTRIBUTIONS

**ECMJ:** Conceptualization, Data curation, Writing – original draft. **JAX:** Data curation, Writing – original draft. **MMS:** Data curation, Writing – original draft. **MCBS:** Conceptualization, Data curation, Funding acquisition , Project administration, Supervision, Writing – original draft.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES

1. UNAIDS. The urgency of now: AIDS at a crossroads. Global AIDS update 2024. Geneva: UNAIDS; 2024. [cited on 2024 Dec 17]. Available from: https://digitallibrary.un.org/record/4086056
2. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing. 2019;48(1):16-31. https://doi.org/10.1093/ageing/afy169
3. Schrock JM. Accelerated aging in people living with HIV: the neuroimmune feedback model. Brain Behav Immun Health. 2024;36:100737. https://doi.org/10.1016/j.bbih.2024.100737
4. Tsakona D, Kapetanakis A, Kyrou D, Cholongitas E, Deutsch M, Malandris K, et al. Mapping frailty in people living with HIV: a nationwide study in Greece. HIV Med. 2023;24(2):170-9. https://doi.org/10.1111/hiv.13356
5. Mu W, Patankar V, Kitchen S, Zhen A. Examining chronic inflammation, immune metabolism, and T cell dysfunction in HIV infection. Viruses. 2024;16(2):219. https://doi.org/10.3390/v16020219
6. Brañas F, Torralba M, Antela A, Vergas J, Ramírez M, Ryan P, et al. Effects of frailty, geriatric syndromes, and comorbidity on mortality and quality of life in older adults with HIV. BMC Geriatr. 2023;23(1):4. https://doi.org/10.1186/s12877-022-03719-8
7. McGee DM, Cotter AG. HIV and fracture: risk, assessment and intervention. HIV Med. 2024;25(5):511-28. https://doi.org/10.1111/hiv.13596
8. Lerner ED, Mahmud IC, Schneider RH. Mortality and osteoporosis in older adults living with HIV/AIDS: a systematic review. Rev Amaz Geriatr Gerontol. 2023;14:16-28. Available from: https://raggfunati.com.br/index.php/ragg/article/view/2
9. Mazzitelli M, Branca IP, Muramatsu T, Carli L, Dusina A, Focà E, et al. FRAX assessment in people ageing with HIV. HIV Med. 2022;23(1):103-8. https://doi.org/10.1111/hiv.13170
10. Lee RC, Wang Z, Heo M, Ross R, Janssen I, Heymsfield SB. Total-body skeletal muscle mass: development and cross-validation of anthropometric prediction models. Am J Clin Nutr. 2000;72(3):796-803. https://doi.org/10.1093/ajcn/72.3.796
11. Alexandre TS, Duarte YA, Santos JL, Wong R, Lebrão ML. Prevalence and associated factors of sarcopenia among elderly in Brazil: findings from the SABE study. J Nutr Health Aging. 2014;18(3):284-90. https://doi.org/10.1007/s12603-013-0413-0
12. Zerbini CA, Szejnfeld VL, Abergaria BH, McCloskey EV, Johansson H, Kanis JA. Incidence of hip fracture in Brazil and the development of a FRAX model. Arch Osteoporos. 2015;10:224. https://doi.org/10.1007/s11657-015-0224-5
13. Zhang HW, Tsai ZR, Chen KT, Lin CC, Lin HY, Chen YM. Enhanced risk of osteoporotic fracture in patients with sarcopenia: a national population-based study in Taiwan. J Pers Med. 2022;12(5):791. https://doi.org/10.3390/jpm12050791
14. Chen S, Xu X, Gong H, Chen R, Guan L, Yan X, et al. Global epidemiological features and impact of osteosarcopenia: a comprehensive meta-analysis and systematic review. J Cachexia Sarcopenia Muscle. 2024;15(1):8-20. https://doi.org/10.1002/jcsm.13392
15. Ahmed M, Mital D, Abubaker NE, Panourgia M, Owles H, Papadaki I, et al. Bone health in people living with HIV/AIDS: an update of where we are and potential future strategies. Microorganisms. 2023;11(3):789. https://doi.org/10.3390/microorganisms11030789
16. Azevedo DC, Hoff LS, Kowalski SC, Andrade CAF, Trevisani VFM, Melo AKG. Risk factors for osteoporotic hip fracture among community-dwelling older adults: a real-world evidence study. Adv Rheumatol. 2024;64:8. https://doi.org/10.1186/s42358-024-00350-6
17. Kanters S, Renaud F, Rangaraj A, Zhang K, Limbrick-Oldfield E, Hughes M, et al. Evidence synthesis evaluating body weight gain among people treating HIV with antiretroviral therapy - a systematic literature review and network meta-analysis. EClinicalMedicine. 2022;48:101412. https://doi.org/10.1016/j.eclinm.2022.101412
18. Huang SG, Lee RP, Yao TK, Wang JH, Wu WT, Yeh KT. Correlation between handgrip strength and bone density and fragility fracture risk among older adults: a cross-sectional study. J Nurs Res. 2025;33(1):e375. https://doi.org/10.1097/jnr.0000000000000656
19. SeyedAlinaghi S, Ghayomzadeh M, Mirzapour P, Maroufi SF, Pashaei Z, Ali Z, et al. A systematic review of sarcopenia prevalence and associated factors in people living with human immunodeficiency virus. J Cachexia Sarcopenia Muscle. 2023;14(3):1168-82. https://doi.org/10.1002/jcsm.13212
20. Huang T, Li C, Chen F, Zhang Y, Liu Y, Li L, et al. Prevalence and risk factors of osteosarcopenia: a systematic review and meta-analysis. BMC Geriatr. 2023;23(1):369. https://doi.org/10.1186/s12877-023-04085-9

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260194

# Determining women's awareness levels of gynecological cancer and their perceptions of breast cancer fatalism

EsmaNur Özen[1], Fatma Ersin[2]*

## SUMMARY

**OBJECTIVE:** This study was conducted to determine women's awareness levels regarding gynecological cancers and their fatalistic perceptions regarding breast cancer.

**METHODS:** This was a descriptive study conducted in a Family Health Centre District between October 2023 and August 2024. The study sample consisted of 333 women. A personal information form, Gynecological Cancer Awareness Scale, and breast cancer fatalism scale were used to collect data.

**RESULTS:** The mean total score on the Gynecological Cancer Awareness Scale for the women participating in the study was 146.02±21.70, and the mean total score on the Breast Cancer Fatalism Scale was 7.40±2.61. Educational status, social security, obtaining information from healthcare professionals about gynecological cancers, and previous gynecological examination status of the participants explain 18.7% of the variance in the Gynecological Cancer Awareness Scale total score average. Additionally, age and educational status explain 13.6% of the variance in the Breast Cancer Fatalism Scale total score average.

**CONCLUSION:** As a result, it can be said that women's awareness levels about gynecological cancers and their perception of breast cancer fatalism are above average. Therefore, education should be provided to increase awareness levels of gynecological cancers and reduce perceptions of breast cancer fatalism.

**KEYWORDS:** Gynecological. Breast. Cancer. Awareness. Nursing.

## INTRODUCTION

Gynecological cancers, defined as malignant diseases of the female reproductive organs, are one of the leading causes of morbidity and mortality in women after breast cancer[1,2]. Gynecological cancers account for 7.24% of all cancer cases[3]. It is estimated that nearly one-fifth of all cancers occurring in women are gynecological cancers[4]. According to Global Cancer Statistics 2022, 6.17% of the 240,013 people newly diagnosed with cancer in our country were diagnosed with gynecological cancer. Of those diagnosed with gynecological cancer, 4.45% died[5]. Awareness among women of the factors that may contribute to gynecological cancers, along with participation in early diagnosis and screening, reduces the risk of death and illness caused by cancer[6]. Awareness levels among women are crucial in preventing gynecological cancers[7]. There are only a limited number of studies examining women's awareness levels regarding gynecological cancers[2,8], and these studies indicate that women's awareness levels are not high.

Although there are many factors affecting women's awareness, one of the important ones is fatalism. Individuals with fatalistic beliefs may think that there is no relationship between cancers and lifestyle and that early diagnosis and screening programs serve no purpose other than diagnosing the disease. This mindset may also lead to less use of early diagnosis services[9].

Breast cancer is one of the cancers with a very high incidence in women. The fatalistic attitude toward breast cancer plays a very important role in the early diagnosis and treatment of breast cancer in women[10]. Individuals who adopt a fatalistic approach may oppose early diagnosis and screening programmes[11]. The literature reports that breast cancer fatalism has a negative impact on early diagnosis and screening behaviours[12]. The literature includes studies with low[13] and high levels of breast cancer fatalism perception[14].

In conclusion, it is evident that awareness levels and fatalistic perceptions play a significant role in the early detection of cancers observed in women. However, the limited number

[1]Harran University, Institute of Health Sciences – Şanlıurfa, Turkey.
[2]Harran University, Faculty of Health Sciences, Department of Public Health Nursing – Şanlıurfa, Turkey.
*Corresponding author: fersin@harran.edu.tr
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on May 02, 2026. Accepted on July 05, 2026.
Scientific Editor: José Maria Soares Júnior

of studies on awareness levels regarding gynecological cancers highlights the need for further research in this area. The literature emphasizes the importance of determining the current situation in order to increase women's awareness levels regarding gynecological cancers. Furthermore, it is thought that women's awareness levels regarding gynecological cancers influence their fatalistic perceptions of breast cancer.

This study was conducted to determine women's awareness levels regarding gynecological cancers and their fatalistic perceptions regarding breast cancer.

### Research questions

What are women's levels of gynecological cancer awareness and breast cancer fatalism perceptions?

What factors influence women's awareness of gynecological cancer and their perception of breast cancer fatalism?

## METHODS

This descriptive study was conducted in the a Family Health Centre Region between October 2023 and August 2024.

In calculating the sample size of the research, the sample calculation method with a known population was used. The population size was taken as 2,497 women between the ages of 30–65, affiliated with a family health center; 95% confidence level (Z=1.96), 5% margin of error (d=0.05), and expected prevalence value p=0.50 were accepted. The minimum number of samples was found to be 333. Women who were not diagnosed with gynecological cancer and/or breast cancer and who had not had a hysterectomy were included in the study. During the sample selection process, the addresses of women who met the inclusion criteria were listed first. The addresses to be visited were then selected randomly using a table. The women were visited at their homes, the purpose of the study was explained to them, and they were given data collection forms. All of the women selected for the sample agreed to participate in the study.

There were no refusals. The data were collected by the researcher visiting the women in their homes and conducting face-to-face interviews lasting 15–20 min. The introductory information form, Gynecological Cancer Awareness Scale (GCAS), and Breast Cancer Fatalism Scale (BCFS) were used to collect the study data.

The Information Form was developed by researchers through a literature review[2,8,15,16].

Gynecological Cancer Awareness Scale (GCAS) consists of 41 items and was developed by Dal and Ertem to determine the level of awareness of gynecological cancers among women aged 20–65 years[16]. The GCAS has four subscales (Awareness of Early Diagnosis and Information on Cancers, Awareness of Gynecological Cancer Risks, Awareness of Prevention of Gynecological Cancers, and Awareness of Routine Check-ups and Perception of Serious Illness in Gynecological Cancers). The Cronbach's alpha value of the scale is 0.944. The score obtained from the scale ranges from a minimum of 41 to a maximum of 205. There is a direct correlation between the scale score and the level of awareness[16]. In our study, the Cronbach's alpha value of the GCAS is 0.91.

Breast Cancer Fatalism Scale (BCFS) is a 15-item scale developed by Powe and revised by Mayo et al to 11 questions[17,18]. Its validity and reliability were established in Turkey by Ersin et al. The scale consists of 11 items, with a "Yes" response scoring 1 point and a "No" response scoring 0 points, resulting in a total score ranging from a minimum of 0 to a maximum of 11. Higher scores on the scale indicate greater fatalism[10]. In our study, the Cronbach's alpha value of the scale was 0.87.

Statistical analysis of data: Data analysis was performed using the SPSS 26.0 software package. The significance level was set at $p<0.05$. Descriptive statistics including frequency, percentage, and mean were used in the analysis of data. Additionally, t-tests for independent groups, Kruskal-Wallis analysis, Pearson correlation analysis, simple linear regression, and multiple linear regression analysis were performed. Skewness and kurtosis coefficients were considered in assessing the suitability of the normal distribution. Skewness and kurtosis coefficients were expected to be in the range of -1 to +1[19].

Ethical principles of the research: Permission was obtained from the Health Directorate (dated 08/12/2023 and numbered E-49781372-799-231234173) and the a University Clinical Research Ethics Committee (dated 13.11.2023 and numbered 2023/21/39) to conduct the research. Informed voluntary consent was obtained from the participants. The study was conducted in accordance with the Helsinki Declaration and the country's ethical standards.

## RESULTS

The study found that 77.2% of women were aged between 30 and 49, 89.8% of women were married, 53.5% of women were illiterate, 51.4% of women had a poor income, 91% of women were unemployed, and 71.8% of women had no social security (Table 1).

About 14.7% of participants stated that they had a gynecological condition, while 32.4% of participants stated that they

had not undergone a gynecological examination. The most frequently cited reasons for not having a gynecological examination were not feeling the need for one (50%), feeling embarrassed about having one (44.7%), and fear of the examination (38.6%). About 74.5% of women stated that they had heard of gynecological cancers, and 69.5% of women stated that they had heard about gynecological cancers from healthcare personnel. In addition, 4.5% of women stated that there was a history of gynecological cancer in their family, and 7.2% of women stated that they knew the symptoms of gynecological cancer.

**Table 1.** Distribution of women's sociodemographic characteristics.

| Variables | Number (n) | Percent (%) |
|---|---|---|
| Age (years) | | |
| Aged 30 to 49 | 257 | 77.2 |
| 50 years old and above | 76 | 22.8 |
| Marital status | | |
| Single | 34 | 10.2 |
| Married | 299 | 89.8 |
| Educational status | | |
| Illiterate | 178 | 53.5 |
| Literate | 61 | 18.3 |
| Primary school | 86 | 25.8 |
| Secondary school and above | 8 | 2.4 |
| Income level | | |
| Poor | 171 | 51.4 |
| Average | 154 | 46.2 |
| Good | 8 | 2.4 |
| Employment status | | |
| Yes | 30 | 9.0 |
| No | 303 | 91.0 |
| Social security status | | |
| Yes | 94 | 28.2 |
| No | 239 | 71.8 |

A significant difference was found in the GCAS total score average among the women participating in the study according to their educational status (KW=10.079, p=0.018), income status (KW=6.966, p=0.031), and social security status (t=3.417, p=0.001). Furthermore, a significant difference was found in the mean total BCFS score according to the women's age (t=-4.934, p=0.000), educational status (KW=35.857, p=0.000), income status (KW=7.653, p=0.022), and social security status (t=2.686, p=0.007).

The mean total GCAS score for women was 146.02±21.70, and the mean total BCFS score was 7.40±2.61. GCAS Early Diagnosis and Information Awareness in Gynecological Cancers sub-dimension was 15.92±2.92, Gynecological Cancer Risks Awareness sub-dimension was 28.4±4.43, Gynecological Cancers Prevention Awareness sub-dimension was 15.51±3.63, and Routine Control and Serious Disease Perception in Gynecological Cancers sub-dimension was 86.85±18.37.

A very weak, negative, statistically significant correlation was found between participants' GCAS scores and their BCFS scores (r=-.222, p=0.000) (Table 2).

Educational status, social security, obtaining information from healthcare professionals about gynecological cancers, and previous gynecological examination status of the participants explain 18.7% of the variance in the GCAS total score average. Additionally, age and educational status explain 13.6% of the variance in the BCFS total score average (Table 3).

## DISCUSSION

The study found that the mean GCAS total score for women was 146.02±21.70. In the literature, GCAS total scores for women range from 148±25.71 to 160.31±22.42[16,20,21]. Considering that the highest possible average score on the scale is 205, it can be said that the level of awareness of gynecological cancers among women in this study is above average. In this study, the mean total score for gynecological cancer awareness among women was found to be 7.40±2.61. In one study in the literature,

**Table 2.** Correlation of mean scores of the Gynecological Cancers Awareness Scale and its subdimensions, and the Breast Cancer Fatalism Scale.

| Gynecological Cancers Awareness Scale and subdimensions average scores | BCFS total score average | |
|---|---|---|
| | r | p |
| Early diagnosis and information awareness sub-dimension in gynecological cancers | -0.265 | 0.000 |
| Awareness of gynecological cancer risks sub-dimension | -0.054 | 0.335 |
| Awareness of prevention of gynecological cancers sub-scale | -0.170 | 0.002 |
| Routine control and perception of serious illness in gynecological cancers sub-dimension | -0.212 | 0.000 |
| GCAS total average score | -0.222 | 0.000 |

BCFS: Breast Cancer Fatalism Scale; GCAS: Gynecological Cancer Awareness Scale.

**Table 3.** Multiple lineage regression analysis between certain characteristics of women and their awareness of gynecological cancers and perceptions regarding fatalism in breast cancer.

| Dependent variables | Model | Independent variables | B | Standard error | Standardized coefficient β | t | p | 95%CI | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | Lower | Upper |
| Breast cancer fatalism | 1 | Age | 0.833 | 0.349 | 0.130 | 2.384 | 0.018 | 0.146 | 1.521 |
| | | Educational status | -0.884 | 0.158 | -0.300 | -5.611 | 0.000 | -1.194 | -0.574 |
| | | Income status | -0.232 | 0.264 | -0.047 | -0.877 | 0.381 | -0.751 | 0.288 |
| | | Social security | 0.617 | 0.317 | 0.103 | 1.946 | 0.053 | -0.007 | 1.241 |
| (R=0.369, $R^2$=0.136, $R^2$ adjusted=0.126, F=12.950; p=0.000, Durbin Watson=1.781) | | | | | | | | | |
| Awareness level of gynecological cancers | 2 | Educational status | 2.805 | 1.225 | 0.118 | 2.290 | 0.023 | 0.395 | 5.214 |
| | | Income status | 2.944 | 2.074 | 0.074 | 1.419 | 0.157 | -1.137 | 7.025 |
| | | Social security | -6.175 | 2.478 | -0.128 | -2.492 | 0.013 | -11.050 | -1.300 |
| | | Obtaining information from healthcare professionals about gynecological cancers | -5.603 | 2.208 | -0.129 | -2.538 | 0.012 | -9.947 | -1.260 |
| | | Previous gynecological examination status | -14.285 | 2.370 | -0.308 | -6.027 | 0.000 | -18.948 | -9.623 |
| (R=0.432; $R^2$=0.187; $R^2$ adjusted=0.175, F=15.044; p=0.000; Durbin Watson=1.614) | | | | | | | | | |

R: correlation coefficient, $R^2$: coefficient determinant, ß: standardized coefficient; CI: confidence interval.

the mean total score for gynecological cancer awareness was 4.69±2.67[21] and 6.00±3.51[15] as stated. In this study, women's above-average fatalistic beliefs may be a hindering factor in implementing health behaviors.

Our study found that women's educational status affects the GCAS total score average. Previous studies also support this finding[8,20], and it is noted that awareness of cervical cancer increases as the level of education rises[21-23]. Considering that education is considered important in raising awareness, this study shows that education explains only a small fraction of awareness. Our study shows that income status affects the average GCAS total score, and the literature indicates that there is a significant relationship between income status and GCAS scores[20]. High income levels among women may have facilitated their access to healthcare services due to their high awareness of gynecological cancer. Women with social security coverage were found to have significantly higher mean GCAS total scores. The literature supports this finding[24]. In this study, having social security may have made it easier for women to communicate with healthcare personnel and increased their awareness.

This study shows that women's age affects the average total BCFS score. A study has indicated that fatalistic attitudes increase with age[25]. As women age, their fatalistic beliefs about breast cancer increase, suggesting that women associate cancer with age. Our study found that women's fatalism regarding breast cancer decreased as their level of education increased, a finding supported by the literature[17]. The fact that highly educated women had lower levels of fatalism suggests that they are more aware of breast cancer. This study shows that women with higher incomes have lower breast cancer fatalism scores. Given that income status facilitates individuals' access to healthcare services, this result obtained from the study is plausible.

According to the regression analysis results, educational status, social security status, prior gynecological examination, and receiving information about gynecological cancers from healthcare personnel explain a low-to-moderate level of variance in gynecological cancer awareness. It can be said that age and education level explain the variance in breast cancer fatalism at a low-to-medium level. The high variance that the models cannot explain suggests that other factors may be effective (cultural beliefs, religious attitudes, health literacy, social support, personal experiences, etc.). Considering that awareness and fatalism are important in the development of positive health behaviors, the results obtained from the study are thought to contribute to the literature and are the expected results. The models show that social security is not significant in terms of fatalism regarding breast cancer. This suggests that, due to women's socio-cultural structure, their sense of privacy is paramount. Furthermore, the perceived ineffectiveness of social security may stem from the collinearity with education. The fatalism scale treats fatalism as a one-dimensional and binary structure for each

item. This binary approach may have caused the variance to appear in this way.

## CONCLUSION AND RECOMMENDATIONS

In this study, it can be said that both the level of awareness of gynecological cancers and the level of fatalism regarding breast cancer were above average. Educational status, social security status, prior gynecological examination, and receiving information about gynecological cancers from healthcare personnel explain 18.7% of the variance in gynecological cancer awareness. Additionally, age and educational status explain 13.6% of the variance in the BCFS total score average.

Health education programs designed to increase women's awareness levels and reduce fatalism tendencies should be planned, implemented, and maintained by nurses in primary healthcare institutions. Furthermore, studies examining women's gynecological cancer awareness and breast cancer fatalism levels should be planned with a larger sample size.

### Limitations

Our study could be conducted with a larger sample size. The results of this study can only be generalized to women aged 30–65 years living in the a family health center area. The study results also show that most of the women are disadvantaged (many are illiterate, lack social security, and are unemployed). Therefore, we cannot generalize the study results to all women. The high rate of illiteracy may have affected the understanding of the data collection tools.

## AUTHOR CONTRIBUTIONS

**EÖ:** Conceptualization, Data curation, Formal Analysis, Writing – original draft. **FE:** Conceptualization, Formal Analysis, Writing – original draft.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Erenoğlu R, Bayraktar E. Awareness levels of married women aged 20-60 years about gynaecological cancer and the affecting factors. J Caring Sci. 2020;13(1):457-69.
2. Gözüyeşil E, Arıöz A, Taş F. Bir aile sağlığı merkezine başvuran kadınların jinekolojik kanser farkındalıklarının değerlendirilmesi. TJFMPC. 2020;14(2):177-85. https://doi.org/10.21763/tjfmpc.730022
3. World Health Organization (WHO). GLOBOCAN 2022. The Global Cancer Observatory. 2024. [cited on 2026 May 05]. Available from: https://gco.iarc.who.int/media/globocan/factsheets/cancers/20-breast-fact-sheet.pdf
4. Eker A, Aslan E. Jinekolojik kanser hastalarında psikososyal yaklaşım. Hemşirelikte Eğitim ve Araştırma Dergisi. 2017;14(4):298-303.
5. World Health Organization (WHO). GLOBOCAN 2022 Turkey. 2024. [cited on 2026 Dec 12]. Available from: https://gco.iarc.who.int/media/globocan/factsheets/populations/792-turkiye-fact-sheet.pdf
6. Öztürk R, Bakir S, Kazankaya F, Paker S, Ertem G. Awareness about gynecologic cancers and related factors among healthy women: a cross-sectional study. Soc Work Public Health. 2021;36(7-8):847-56. https://doi.org/10.1080/19371918.2021.1965936
7. Ersin F, Kahraman S, Havlioğlu S. Kadınların jinekolojik kanserler farkındalık durumları ve etkileyen faktörler: Araştırma Makalesi. European Health Sci J. 2024;2(2):28-34. https://doi.org/10.5281/zenodo.13149326
8. Atlas B, Güneri SE. Kadınların jinekolojik kanserlerle ilgili farkındalığı ve farkındalığı etkileyen faktörler. İzmir Kâtip Çelebi Üniversitesi Sağlık Bilimleri Fakültesi Dergisi. 2022;7(1):77-85.
9. Lutalo T, Namusoke F. Health information sources and contraceptive use among women. Int J Public Health. 2023;68:1-9.
10. Ersin F, Capik C, Kissal A, Aydogdu NG, Beser A. Breast cancer fatalism scale: a avalidity and reliability study in Turkey. IJCS. 2018;11(2):783.
11. Ülger E, Alacacıoğlu A, Gülseren AŞ, Zencir G, Demir L, Tarhan MO. Kanserde psikososyal sorunlar ve psikososyal onkolojinin önemi. Dokuz Eylül Üniversitesi Tıp Fakültesi Dergisi. 2014;28(2):85-92.
12. Ramírez AS. Fatalism and cancer risk knowledge among a sample of highly acculturated Latinas. J Cancer Educ. 2014;29(1):50-5. https://doi.org/10.1007/s13187-013-0541-6
13. Altintas HK, Ayyildiz TK, Veren F, Topan AK. The effect of breast cancer fatalism on breast cancer awareness among Turkish women. J Relig Health. 2017;56(5):1537-52. https://doi.org/10.1007/s10943-016-0326-4
14. Vrinten C, Wardle J, Marlow LA. Cancer fear and fatalism among ethnic minority women in the United Kingdom. Br J Cancer. 2016;114(5):597-604. https://doi.org/10.1038/bjc.2016.15
15. Aksu B, Ersin F. The effect of breast cancer fatalism and barrier perceptions of female seasonal agricultural workers on their breast cancer early detection behaviors. IJCS. 2023;16(1):386-95.
16. Dal NA, Ertem G. Jinekolojik kanserler farkındalık ölçeği geliştirme çalışması. İnsan ve Toplum Bilimleri Araştırmaları Dergisi. 2017;6(5):2351-67.
17. Powe BD, Hamilton J, Brooks P. Perceptions of cancer fatalism and cancer knowledge: a comparison of older and younger African American women. J Psychosoc Oncol. 2006;24(4):1-13. https://doi.org/10.1300/J077v24n04_01
18. Mayo RM, Ureda JR, Parker VG. Importance of fatalism in understanding mammography screening in rural elderly women.

J Women Aging. 2001;13(1):57-72. https://doi.org/10.1300/J074v13n01_05

19. George D, Mallery P. SPSS for windows: a simple guide and reference 10.0 update. Allyn & Bacon Company; 2001.
20. Çelebi N. Kadınların jinekolojik kanserlere yönelik farkındalıkları ve erken tanı yöntemlerine yönelik davranışlarının değerlendirilmesi. [Yüksek Lisans Tezi, Atatürk Üniversitesi, Sağlık Bilimleri Enstitüsü]. 2021. Available from: https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp
21. Tuncer SK, Karakurt P. Kadınların jinekolojik kanserler ile ilgili farkındalık düzeyinin artmasında sağlık okuryazarlığının etkisi üzerine bir araştırma. Mersin Üniversitesi Tıp Fakültesi Lokman Hekim Tıp Tarihi ve Folklorik Tıp Dergisi. 2023;13(1):196-206. https://orcid.org/0000-0003-0330-9807
22. Han B, Baysal HY. Relationship of women's breast cancer fatalism perceptions with health beliefs and early diagnosis behaviors. Adv Women's Stud. 2024;6(2):38-47. https://doi.org/10.51621/aws.1553428
23. Silva AHW, Samarawickrema N, Kasturiratne A, Skinner SR, Wickremasinghe AR, Garland SM. Awareness of human papillomavirus, cervical cancer and its prevention among primigravid antenatal clinic attendees in a tertiary care hospital in Sri Lanka: a cross-sectional study. Sex Health. 2019;16(3):212-7. https://doi.org/10.1071/SH18080
24. Gyamfua AA, Nkrumah I, Ibitoye BM, Agyemang BA, Ofosu ES, Tsoka-Gwegweni JM, et al. The level of knowledge and associated socio-demographic factors on cervical cancer among women: a cross-sectional study at Kenyase Bosore community, Ghana. Pan Afr Med J. 2019;34:44. https://doi.org/10.11604/pamj.2019.34.44.19471
25. Jamieson PE, Romer D. Unrealistic fatalism in U.S. youth ages 14 to 22: prevalence and characteristics. J Adolesc Health. 2008;42(2):154-60. https://doi.org/10.1016/j.jadohealth.2007.07.010

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260286

# Evaluation of serum soluble triggering receptor expressed on myeloid cells-1 in polycystic ovary syndrome

Taylan Onat[1]*, Senem Arda Duz[2], Erdinc Saridogan[2], Selami Gunal[3], Gamze Nur Cimilli Senocak[4], Gorkem Tuncay[2], Abdullah Karaer[2]

## SUMMARY

**OBJECTIVE:** This study aims to investigate the clinical significance of serum soluble triggering receptor expressed on myeloid cells-1 levels in patients with polycystic ovary syndrome.

**METHODS:** This case-control study included 40 women with polycystic ovary syndrome and 40 healthy control women. Serum soluble triggering receptor expressed on myeloid cells-1 and high-sensitive C-reactive protein levels were measured via commercially available enzyme-linked immunosorbent assay, according to the manufacturer's instructions.

**RESULTS:** Soluble triggering receptor expressed on myeloid cells-1 levels were statistically significantly higher in women with polycystic ovary syndrome than controls ($p<0.001$). In the logistic regression analysis, an increase in serum soluble triggering receptor expressed on myeloid cells-1 level demonstrated a significant association with polycystic ovary syndrome diagnosis ($p=0.011$). In the receiver operating characteristic curve analysis, the area under the curve (AUC) for soluble triggering receptor expressed on myeloid cells-1 was found to be 0.800 ($p<0.001$). Setting the cut-off value at 45.22 pg/mL, soluble triggering receptor expressed on myeloid cells-1 demonstrated a sensitivity of 87.5% and a specificity of 65.0%. In the logistic regression analysis that incorporated homeostatic model assessment for insulin resistance, anti-müllerian hormone, and soluble triggering receptor expressed on myeloid cells-1, a 1 pg/mL rise in soluble triggering receptor expressed on myeloid cells-1 was associated with a 1.07-fold increase in the odds of polycystic ovary syndrome, a 3.10-fold increase in the odds due to a 1-point rise in homeostatic model assessment for insulin resistance, and a 3.24-fold increase in the odds for every 1 ng/mL rise in anti-müllerian hormone. The model incorporating anti-müllerian hormone, homeostatic model assessment for insulin resistance, and soluble triggering receptor expressed on myeloid cells-1 predicted polycystic ovary syndrome with 83.8% accuracy, 90.0% specificity, and 77.5% sensitivity.

**CONCLUSION:** Soluble triggering receptor expressed on myeloid cells-1 may serve as an indicator of the chronic inflammatory process in polycystic ovary syndrome and could be a complementary marker rather than a diagnostic marker for polycystic ovary syndrome.

**KEYWORDS:** Chronic inflammation. Hyperandrogenism. Insulin resistance. Polycystic ovary syndrome. STREM-1.

## INTRODUCTION

Polycystic ovary syndrome (PCOS) is a common endocrinological disorder with varied prevalence depending on the criteria used[1]. The typical clinical presentation frequently includes oligo/amenorrhea, hirsutism, acne, androgenic alopecia, and obesity[2]. The phenotypes of PCOS include a variety of findings and symptoms. It has been found that different phenotypes in PCOS differ in terms of metabolic risks[3]. In addition to the Rotterdam criteria, the National Institutes of Health (NIH) and Androgen Excess-PCOS (AE-PCOS) criteria are also used for diagnosing PCOS. The Rotterdam criteria are used more frequently because they are based on the current recommendations of the 2023 international evidence-based guidelines for the assessment and management of PCOS[2]. Infertility, type 2 diabetes mellitus, dyslipidemia, cardiovascular diseases, cerebrovascular events, and endometrial cancer may occur in the long term[2]. Although PCOS was defined in 1935 and its pathogenesis has been investigated by many studies, its etiopathogenesis remains poorly understood. According to ethnicity and age, the clinical presentation may vary considerably, making PCOS a highly heterogeneous condition[2]. While menstrual irregularities such as oligo/amenorrhea are hallmark features of PCOS, distinguishing these from physiological anovulation during adolescence remains challenging. A recent systematic review and meta-analysis by Zuchelo et al. highlighted the overlap between menstrual patterns in PCOS

[1]University of Health Sciences, Sincan Training and Research Hospital – Ankara, Türkiye.
[2]Inonu University, School of Medicine, Department of Obstetrics and Gynecology – Malatya, Türkiye.
[3]Inonu University, School of Pharmacy, Department of Pharmaceutical Microbiology – Malatya, Türkiye.
[4]Ataturk University, School of Medicine, Department of Obstetrics and Gynecology – Erzurum, Türkiye.
*Corresponding author: onat.taylan@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on January 09, 2026. Accepted on July 10, 2026.
Scientific Editor: José Maria Soares Júnior

and hypothalamic–pituitary–ovarian axis immaturity in adolescents, emphasizing the need for careful diagnostic evaluation beyond menstrual history alone[4].

Evidence suggests that androgens contribute to the inflammatory milieu in patients with PCOS[5]. Furthermore, visceral adipose tissue has been implicated in the chronic inflammatory state observed in patients with PCOS[6]. In addition, several inflammatory markers, including chemokines and cytokines, have been identified in the serum of women with PCOS[5].

The triggering receptor expressed on myeloid cells 1 (TREM-1) is also associated with the inflammatory process[7]. TREM-1 promotes secretion of interleukin-1β (IL-1β), tumor necrosis factor-α (TNF-α), and monocyte chemotactic protein-1 and initiates the production of some inflammatory cytokines by induction of Toll-like receptor and nucleotide-binding oligomerization domain (NOD)-like receptor pathways[8]. The aim of this study was to investigate the clinical significance of serum soluble triggering receptor expressed on myeloid cells-1 (sTREM-1) levels in patients with PCOS. To the best of our knowledge, this is the first study to evaluate sTREM-1 as an inflammatory marker in patients with PCOS.

## METHODS

This study was designed as a case-control study and was conducted in a tertiary medical centre. The study was performed in accordance with the ethical principles outlined in the Helsinki Declaration. The study period was from July 2022 to January 2023. Atatürk University Faculty of Medicine Ethics Committee for Clinical Research approval was obtained with the number (30.06.2022-6-23). All participants provided written informed consent.

Based on a study conducted by Rudnicka et al.[9], with a statistical power of 80% and an alpha of 20%, a total of 37 participants per group was needed for the study. We decided on 40 individuals in each group, assuming a dropout rate of 10%. The study included women diagnosed with PCOS aged 18 to 30 years. The diagnosis of PCOS was made using the Rotterdam criteria. Exclusion criteria included the presence of chronic systemic diseases (such as diabetes mellitus), acute infections, and use of oral contraceptives and/or insulin sensitizers within the preceding three months. In addition, the use of medications that affect inflammatory markers (metformin, NSAIDs, etc.) and endocrine disorders such as Cushing's syndrome, congenital adrenal hyperplasia, and thyroid disease were also used as exclusion criteria. Their body mass indexes (BMI), waist–hip ratios (WHR), and Ferriman-Gallwey scores (FGS) were recorded. Fasting venous blood samples were collected during the early follicular phase, following an overnight fast of at least 8 h.

Serum samples acquired during the fasting period from patients with PCOS were obtained within the 2nd–5th day of menstruation and subsequently preserved at -80°C for subsequent analysis. Serum sTREM-1 (Elapscience, Catalog No: E-EL-H1596, USA) and hs-CRP levels were measured using commercially available enzyme-linked immunosorbent assay (ELISA) kits, according to the manufacturer's instructions. After the possibility of pregnancy was ruled out in the group of patients with oligo-amenorrhea, blood samples were collected between the 2nd and 5th days of progesterone-induced menstruation.

Statistical analyses were performed using SPSS Statistics software, version 22. The Kolmogorov-Smirnov test was used to assess the normality of continuous variables. Depending on whether the quantitative data were normally distributed or not, the Student's t-test or Mann-Whitney U test was used. Categorical variables were compared by the chi-square test. To evaluate the associations between numerical data, either the Pearson or Spearman Rho test was employed. The impact of sTREM-1 levels and covariates on PCOS diagnosis was assessed using multivariate binary logistic regression. To avoid multicollinearity, several variables that were significant in the univariate analysis were deliberately excluded from the final model because they were closely related to the covariates that were retained. Specifically, fasting insulin was excluded because it is a component of HOMA-IR, while waist–hip ratio and testosterone were excluded because they closely reflect the metabolic and androgenic status that is already captured by HOMA-IR and AMH. Receiver operating characteristic (ROC) curve analysis was performed to evaluate the predictive performance of sTREM-1 levels for PCOS diagnosis. Statistical significance was defined as $p<0.05$.

## RESULTS

Table 1 presents the demographic characteristics, biochemical values, and sTREM-1 levels of the study participants. The two groups did not differ significantly with respect to BMI or age. Statistically significant differences were observed in waist-to-hip ratio (WHR) and FGS between the groups ($p<0.05$). sTREM-1 levels were significantly higher in women with PCOS than in controls ($p<0.001$). Additionally, total testosterone, free testosterone, thyroid-stimulating hormone (TSH), anti-mullerian hormone (AMH), and hsCRP levels were found to be statistically significant between the two groups ($p<0.05$) (Table 1). Analysis of the relationships among independent variables

revealed a significant positive correlation between sTREM-1 and hsCRP (r=0.471, p<0.001) and AMH (r=0.274, p=0.014), although the latter reflects only a weak association. Additionally, a statistically significant correlation was found between AMH and hsCRP (r=0.251, p=0.024) and Ferriman-Gallwey score (FGS) (r=0.344, p=0.010) (Table 2).

ROC curve analysis demonstrated AUC values of 0.800 for sTREM-1 and 0.785 for AMH (Figure 1). Setting the cut-off value at 2.81 ng/mL, AMH demonstrated sensitivity of 77.5%, specificity of 70.0%, and accuracy of 73.75%. For sTREM-1, at the cut-off value determined as 45.22 pg/mL, specificity was detected to be 65.0%, sensitivity 87.5%, and accuracy 76.25%.

In the logistic regression analysis that incorporated HOMA-IR, AMH, and sTREM-1, a 1 pg/mL rise in sTREM-1 was associated with a 1.07-fold increase in the odds of PCOS, a 3.10-fold increase in the odds due to a 1-point rise in HOMA-IR, and a 3.24-fold increase in the odds for every 1 ng/mL rise in AMH. The model incorporating AMH, HOMA-IR, and sTREM-1 predicted PCOS with 83.8% accuracy, 90.0% specificity, and 77.5% sensitivity.

## DISCUSSION

This study is the first to demonstrate an association between sTREM-1 levels and PCOS. The fact that sTREM-1 shows a moderate positive correlation with hsCRP supports the hypothesis that sTREM-1 may serve as a marker of chronic low-grade inflammation, which plays a role in the pathogenesis of PCOS. Notably, the PCOS group exhibited significantly higher levels of sTREM-1 and AMH. Furthermore, sTREM-1, AMH,

**Table 1.** Clinical and laboratory characteristics of the women with or without polycystic ovary syndrome.

| | PCOS (n: 40) | Control (n: 40) | P-value |
|---|---|---|---|
| Age (years) | 20.5 (20.5–23) | 22 (20–23) | 0.100µ |
| BMI (kg/m$^2$) | 22.2 (20.1–26.6) | 21.2 (19.6–23.3) | 0.132µ |
| WHR | 0.797±0.072 | 0.757±0.068 | **0.014*** |
| FGS | 11 (9.75–19.3) | 0 (0–5) | **<0.001**µ |
| HOMA-IR | 2.13 (1.67–3.1) | 1.48 (1.15–2.0) | **<0.001**µ |
| AMH (ng/mL) | 3.66 (2.88–4.19) | 2.5 (1.96–3.08) | **<0.001**µ |
| FSH (mIU/mL) | 5.57±1.31 | 5.64±1.42 | 0.817* |
| LH (mIU/mL) | 6.46 (4.04–9.31) | 5.61 (4.66–5.99) | 0.130µ |
| Estradiol (pg/mL) | 40.2 (31.2–56.0) | 40.7 (38.7–51.4) | 0.301µ |
| Prolactin (ng/mL) | 14.2 (9.04–19.1) | 12.6 (9.73–15.1) | 0.229µ |
| TSH (mIU/L) | 2.39 (1.89–3.14) | 1.92 (1.49–2.43) | **0.015**µ |
| $fT_4$ (ng/mL) | 1.15 (1.02–1.22) | 1.08 (0.96–1.16) | 0.050µ |
| DHEA-S (µg/dL) | 230 (174–275) | 199 (166–231) | 0.138µ |
| Total testosterone (ng/dL) | 31.9 (25.1–53.7) | 24.0 (20.0–28.8) | **<0.001**µ |
| Free testosterone (ng/dL) | 2.19 (1.84–2.72) | 1.83 (1.57–2.06) | **0.004**µ |
| Fasting glucose (mg/dL) | 87.5 (83–91) | 86 (81.5–90.5) | 0.424µ |
| Fasting insulin (mIU/mL) | 9.66 (7.51–14.8) | 6.94 (5.56–9.5) | **0.001**µ |
| TC (mg/dL) | 172 (157–186) | 166 (147–179) | 0.353µ |
| LDL-C (mg/dL) | 101 (86.5–115) | 91.5 (85.5–104) | 0.184µ |
| HDL-C (mg/dL) | 52.8±10.8 | 55.1±10.3 | 0.340* |
| TG (mg/dL) | 82 (70.8–109) | 87 (63.8–107) | 0.736µ |
| hsCRP (mg/L) | 2.25±0.887 | 1.81±0.663 | **0.013*** |
| sTREM-1 (pg/mL) | 61.4±21.3 | 38.2±20.3 | **<0.001*** |

Data are presented as median (25th–75th percentiles) and mean±standard deviation. The Mann-Whitney U (µ) and Student's t-tests (*) were used. BMI: body mass index; WHR: waist–hip ratio; FGS: Ferriman-Gallwey score; HOMA-IR: homeostatic model assessment for insulin resistance; AMH: anti-müllerian hormone; FSH: follicle-stimulating hormone; LH: luteinizing hormone; TSH: thyroid-stimulating hormone; $fT_4$: free thyroxine; DHEA-S: dehydroepiandrosterone sulfate; TC: total cholesterol; LDL-C: low-density lipoprotein cholesterol; HDL-C: high-density lipoprotein cholesterol; TG: triglyceride; hsCRP: high-sensitive C-reactive protein; sTREM-1: soluble form of triggering receptor expressed on myeloid cells; PCOS: polycystic ovary syndrome. Bold values indicate statistical significance (p<0.05).

and HOMA-IR were identified as factors significantly associated with PCOS.

The relationship between AMH, IR, and inflammation plays a significant role in the pathophysiology of PCOS. IR and the resulting hyperinsulinemia are considered key drivers of PCOS pathophysiology, contributing to hyperandrogenism and reproductive dysfunction[10]. Unlike other target organs, such as the liver and skeletal muscle, the ovaries maintain insulin sensitivity through the phosphatidylinositol 3-kinase-independent insulin signaling pathway[11]. Insulin-induced increased androgen synthesis by theca cells increases the number of pre-antral follicles, which causes a rise in AMH, disrupting follicular maturation. Elevated AMH, in conjunction with insulin resistance, has been demonstrated to suppress the conversion of androgens to estrogen in granulosa cells by inducing aromatase suppression. Tumor necrosis factor alpha (TNF-α) can cause insulin resistance by inhibiting insulin receptor tyrosine kinase in target organs[12]. There are studies suggesting that TREM1 may regulate the release of proinflammatory cytokines such as TNF-α and IL-8[13]. The relationship between hyperandrogenism, insulin resistance (IR), and chronic low-grade inflammation in PCOS is well-established. In this study, significant differences were observed between the two groups in the clinical and biochemical findings of hyperandrogenism, HOMA-IR values, and inflammation markers. However, no correlation was found among these parameters. The potential causes of this phenomenon may include the heterogeneous nature of PCOS and the limited sample size of the study. Alternatively, sTREM-1 may reflect a systemic inflammatory state that is partly independent of the metabolic and androgenic features of PCOS, or

**Table 2.** The relationships between independent variables.

| | | sTREM-1 | HOMA-IR | WHR | FGS | Total testosterone | hsCRP |
|---|---|---|---|---|---|---|---|
| HOMA-IR | Correlation coefficient | 0.143 | – | | | | |
| | p-value | 0.206 | – | | | | |
| WHR | Correlation coefficient | 0.122 | 0.130 | – | | | |
| | p-value | 0.281 | 0.250 | – | | | |
| FGS | Correlation coefficient | 0.289 | 0.420 | 0.180 | – | | |
| | p-value | **0.009** | **<0.001** | 0.109 | – | | |
| Total testosterone | Correlation coefficient | 0.147 | 0.027 | 0.215 | 0.350 | – | |
| | p-value | 0.194 | 0.814 | 0.055 | **0.001** | – | |
| hsCRP | Correlation coefficient | 0.373 | 0.021 | 0.183 | 0.159 | 0.313 | – |
| | p-value | **<0.001** | 0.856 | 0.104 | 0.160 | **0.005** | – |
| AMH | Correlation coefficient | 0.274 | 0.135 | 0.084 | 0.186 | 0.199 | 0.251 |
| | p-value | **0.014** | 0.233 | 0.458 | 0.099 | 0.077 | **0.024** |

Correlation analyses of independent variables were performed using Spearman's Rho tests. sTREM-1: soluble form of triggering receptor expressed on myeloid cells; HOMA-IR: homeostatic model assessment for insulin resistance; WHR: waist–hip ratio; FGS: Ferriman-Gallwey score; hsCRP: high-sensitive C-reactive protein; AMH: anti-müllerian hormone. Bold values indicate statistical significance (p<0.05).

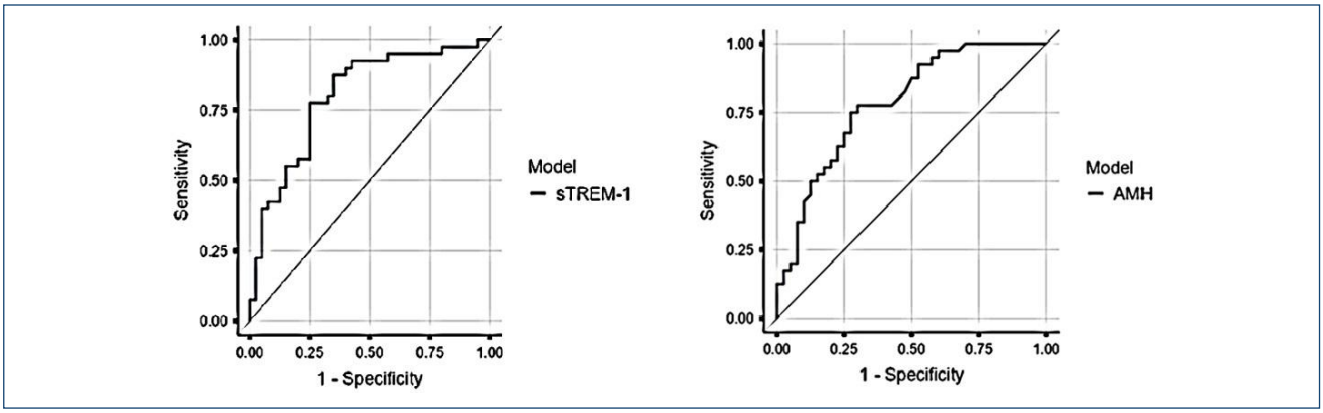



**Figure 1.** Receiver operating characteristiccurve analysis of soluble triggering receptor expressed on myeloid cells-1 and anti-müllerian hormone.

it may be influenced by unmeasured factors such as adiposity, dietary habits, and physical activity, which were not assessed in the present study.

Diagnosing PCOS during adolescence is challenging because oligo-amenorrhea is frequently observed during this period, as the hypothalamic–pituitary axis has not yet fully matured. There is currently no consensus on this issue. In their meta-analysis, Zuchelo et al. found that adolescents with PCOS had a higher BMI compared to adolescents with oligo-amenorrhea and that amenorrhea is a significant finding for PCOS[4]. The results of our study support the notion that sTREM could be a candidate marker for use as a complementary marker in suspected cases.

Inflammation is a normal immunological response to injury or infection; however, when persistent, it may cause tissue damage and contribute to the development of chronic conditions such as PCOS. Recent evidence suggests that PCOS may be related to chronic low-grade inflammation[14]. Inflammation is a diverse biological response that is activated in response to cellular injury and is involved in many physiological processes. The chronic inflammatory state observed in PCOS may have implications beyond metabolic and reproductive dysfunction. Baracat et al. demonstrated that hormonal and metabolic factors influence progesterone action on the endometrium in women with PCOS, potentially contributing to the endometrial dysfunction and increased cancer risk associated with this condition[15]. While our study focused on sTREM-1 as a marker of systemic inflammation, future research should explore whether elevated sTREM-1 levels correlate with endometrial inflammatory markers and whether this contributes to the endometrial abnormalities observed in PCOS. One of the theories is that the ovaries of women with PCOS are subject to chronic low-level oxidative stress, which can trigger an inflammatory response[16]. Oxidative stress is thought to result from factors such as insulin resistance and elevated androgen levels, which are commonly observed in women with PCOS[16].

Elevated hs-CRP concentrations have been associated with increased cardiovascular risk, which is of particular concern in women with PCOS, who are already at elevated risk for cardiovascular disease[17]. Several studies have investigated the role of hs-CRP in PCOS[18]. These studies have found that women with PCOS have higher levels of hs-CRP than controls[18]. Furthermore, research suggests that inflammatory markers may not be raised in women with non-obese PCOS[19]. These discrepancies may be attributed to differences in study design, patient populations, and other methodological factors.

AMH is a member of the transforming growth factor beta family, which has been used to assess ovarian reserve until recent years. However, its diagnostic performance, particularly in PCOS, has now begun to be included in guidelines[2]. When women with ovulatory cycles are compared to women with PCOS, it has been found that women with PCOS have higher AMH levels. Barghi et al.[20] discovered in their meta-analysis that AMH levels are elevated in women with PCOS and diminish with advancing age and increasing BMI. In the study, we used two groups matched for age and BMI to isolate the effects of these factors on AMH. The results of our study align with existing literature on the subject, which has demonstrated that AMH levels are elevated in individuals diagnosed with PCOS.

sTREM-1, a biomarker of inflammation, plays a crucial role in triggering the innate immune response. This cell surface molecule is expressed by both neutrophils and monocytes, which are actively involved in fueling the propagation of the inflammatory response[21]. TREM-1 levels are influenced by both infectious and inflammatory diseases[21-24]. Evidence indicates that sTREM-1 promotes the generation of pro-inflammatory cytokines, including TNF-$\alpha$[22]. Ferat-Osorio et al. noted that sTREM-1 reflects an inflammatory state rather than an infectious one[23]. TREM-1 expression was shown to be greater in diabetes-obese individuals compared to non-diabetic obese patients, implying a link between insulin resistance and TREM-1[24]. Notably, sTREM-1 represents a promising therapeutic target for pharmacological intervention[22]. In this study, sTREM-1 levels were found to be significantly higher in women with PCOS. Notably, this difference persisted despite the PCOS and control groups being matched for BMI, suggesting that the elevation in sTREM-1 reflects an intrinsic PCOS-related inflammatory component rather than merely an effect of adiposity.

Although insulin resistance is not a diagnostic criterion for PCOS, it is frequently observed even in lean women with PCOS. Lim et al. conducted a meta-analysis that revealed the detrimental effect of obesity on metabolic and reproductive outcomes in women with PCOS[25]. Guo et al. found that the clinical and biochemical signs of PCOS have been altered with age[26]. According to that study, symptoms of hyperandrogenism predominate at younger ages, while metabolic symptoms become more prominent with advancing age[26].

There are some limitations of this study. Because this study is a case-control study, a cause-and-effect relationship could not be established. Additional limitations include the small sample size, the inability to stratify participants according to PCOS subgroups, and the use of a single blood sample per participant. In addition, factors such as excess weight, physical activity, and dietary habits, which we did not assess in this study, represent a significant limitation. Different PCOS phenotypes exhibit varying degrees of metabolic dysfunction and inflammatory profiles. Soares-Jr et al. demonstrated that PCOS phenotypes

influence the development of metabolic syndrome over a six-year follow-up, with phenotype A (hyperandrogenism+oligo/amenorrhea+polycystic ovary morphology) showing the highest metabolic risk[3]. The small sample size may have limited statistical power to detect modest correlations between sTREM-1 and metabolic parameters such as HOMA-IR or clinical markers of hyperandrogenism (FGS); these null findings should therefore be interpreted with caution and do not preclude a true association. On the other hand, this study is the first to examine the relationship between PCOS and sTREM-1.

In conclusion, sTREM-1 appears to be associated with the chronic inflammatory process in PCOS and may contribute to our understanding of the role of inflammation in its pathogenesis. sTREM-1 may serve as an indicator of the chronic inflammatory process in PCOS and could be a complementary marker rather than a diagnostic marker for PCOS. Given the heterogeneous nature of PCOS, these findings need to be validated by studies conducted in larger, well-defined cohorts that include both lean and obese women from different PCOS subgroups. In addition, future studies could evaluate sTREM-1 levels in adolescents and in populations receiving PCOS treatment.

## AUTHORS' CONTRIBUTIONS

**TO:** Conceptualization, Data curation, Methodology, Writing – original draft. **SAD:** Data curation, Writing – original draft. ES: Investigation, Writing – original draft. **SG:** Formal Analysis. **GNCS**: Data curation, Investigation. **GT:** Methodology, Project administration, Writing – review & editing. **AK:** Conceptualization, Project administration, Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Chiaffarino F, Cipriani S, Dalmartello M, Ricci E, Esposito G, Fedele F, et al. Prevalence of polycystic ovary syndrome in European countries and USA: a systematic review and meta-analysis. Eur J Obstet Gynecol Reprod Biol. 2022;279:159-70. https://doi.org/10.1016/j.ejogrb.2022.10.020
2. Teede HJ, Tay CT, Laven JJE, Dokras A, Moran LJ, Piltonen TT, et al. Recommendations from the 2023 international evidence-based guideline for the assessment and management of polycystic ovary syndrome. Eur J Endocrinol. 2023;189(2):G43-64. https://doi.org/10.1093/ejendo/lvad096
3. Soares-Jr JM, Hayashida SAY, Marcondes JAM, Maciel GAR, Barcellos CRG, Maffazioli GN, et al. Influence of phenotypes on the metabolic syndrome of women with polycystic ovary syndrome over a six-year follow-up in Brazil. Biomedicines. 2023;11(12):3262. https://doi.org/10.3390/biomedicines11123262
4. Zuchelo LTS, Alves MS, Baracat EC, Sorpreso ICE, Soares JM. Menstrual pattern in polycystic ovary syndrome and hypothalamic-pituitary-ovarian axis immaturity in adolescents: a systematic review and meta-analysis. Gynecol Endocrinol. 2024;40(1):2360077. https://doi.org/10.1080/09513590.2024.2360077
5. Abraham Gnanadass S, Divakar Prabhu Y, Valsala Gopalakrishnan A. Association of metabolic and inflammatory markers with polycystic ovarian syndrome (PCOS): an update. Arch Gynecol Obstet. 2021;303(3):631-43. https://doi.org/10.1007/s00404-020-05951-2
6. Patel S. Polycystic ovary syndrome (PCOS), an inflammatory, systemic, lifestyle endocrinopathy. J Steroid Biochem Mol Biol. 2018;182:27-36. https://doi.org/10.1016/j.jsbmb.2018.04.008
7. Trivedi N, Bhosale JD, Pant A, Suryawanshi SP, Tiwari P, Abel PW, et al. Triggering receptor expressed on myeloid cells-1 (TREM-1) in inflammation and disease: mechanisms, therapeutic potential, and future directions. Int J Mol Sci. 2025;26(21):10386. https://doi.org/10.3390/ijms262110386
8. Prüfer S, Weber M, Sasca D, Teschner D, Wölfel C, Stein P, et al. Distinct signaling cascades of TREM-1, TLR and NLR in neutrophils and monocytic cells. J Innate Immun. 2014;6(3):339-52. https://doi.org/10.1159/000355892
9. Rudnicka E, Kunicki M, Suchta K, Machura P, Grymowicz M, Smolarczyk R. Inflammatory markers in women with polycystic ovary syndrome. Biomed Res Int. 2020;2020:4092470. https://doi.org/10.1155/2020/4092470
10. Kisiała M, Mazepa W, Bauer D, Tabaka P, Gas M, Sowiński A, et al. Molecular mechanisms of insulin resistance and altered carbohydrate metabolism in PCOS: a scoping review. Front Endocrinol (Lausanne). 2026;17:1810805. https://doi.org/10.3389/fendo.2026.1810805
11. Poretsky L, Seto-Young D, Shrestha A, Dhillon S, Mirjany M, Liu HC, et al. Phosphatidyl-inositol-3 kinase-independent insulin action pathway(s) in the human ovary. J Clin Endocrinol Metab. 2001;86(7):3115-9. https://doi.org/10.1210/jcem.86.7.7617
12. Mittal R, Prasad K, Lemos JRN, Arevalo G, Hirani K. Unveiling gestational diabetes: an overview of pathophysiology and management. Int J Mol Sci. 2025;26(5):2320. https://doi.org/10.3390/ijms26052320
13. Wang YS, Li XJ, Zhao WO. TREM-1 is a positive regulator of TNF-α and IL-8 production in U937 foam cells. Bosn J Basic Med Sci. 2012;12(2):94-101. https://doi.org/10.17305/bjbms.2012.2503
14. Rostamtabar M, Esmaeilzadeh S, Tourani M, Rahmani A, Baee M, Shirafkan F, et al. Pathophysiological roles of chronic low-grade inflammation mediators in polycystic ovary syndrome. J Cell Physiol. 2021;236(2):824-38. https://doi.org/10.1002/jcp.29912
15. Baracat MCP, Baracat EC, Simões RS, Simões MJ, Maciel GAR, Azziz R, et al. Hormonal and metabolic factors influence the action of progesterone on the endometrium of women with polycystic ovary syndrome. Diagnostics (Basel). 2023;13(3):382. https://doi.org/10.3390/diagnostics13030382

16. Zhao W, Zhou J, Song Y, He M, Zhu X, Yu B, et al. Polycystic ovary syndrome revisited: novel insights and updates. Int J Med Sci. 2026;23(1):271-82. https://doi.org/10.7150/ijms.119968

17. Karaer A, Cavkaytar S, Mert I, Buyukkagnici U, Batioglu S. Cardiovascular risk factors in polycystic ovary syndrome. J Obstet Gynaecol. 2010;30(4):387-92. https://doi.org/10.3109/01443611003763408

18. Mazloomi S, Sheikh N, Sanoee Farimani M, Pilehvari S. Association of Prx4, total oxidant status, and inflammatory factors with insulin resistance in polycystic ovary syndrome. Int J Endocrinol. 2021;2021:9949753. https://doi.org/10.1155/2021/9949753

19. Moin ASM, Sathyapalan T, Atkin SL, Butler AE. Inflammatory markers in non-obese women with polycystic ovary syndrome are not elevated and show no correlation with vitamin D metabolites. Nutrients. 2022;14(17):3540. https://doi.org/10.3390/nu14173540

20. Barghi M, Heidari Z, Haghighatdoost F, Feizi A, Hashemipour M. New insights into the relationship of antimüllerian hormone with polycystic ovary syndrome and its diagnostic accuracy: an updated and extended meta-analysis using a marginal beta-binomial model. Am J Obstet Gynecol. 2025;232(2):164-87.e31. https://doi.org/10.1016/j.ajog.2024.10.004

21. Tang Q, He M, Zhang S, Zhang J, Yang L, Shi H. The diagnostic value of triggering receptor expressed on myeloid cells-1 in post-traumatic bacterial endophthalmitis. Invest Ophthalmol Vis Sci. 2023;64(5):4. https://doi.org/10.1167/iovs.64.5.4

22. Brustolin B, Touly N, Maillefer M, Parisot L, Pillo E, Derive M, et al. Triggering receptor expressed on myeloid cells-1 deletion in mice attenuates high-fat diet-induced obesity. Front Endocrinol (Lausanne). 2023;13:983827. https://doi.org/10.3389/fendo.2022.983827

23. Ferat-Osorio E, Wong-Baeza I, Esquivel-Callejas N, Figueroa-Figueroa S, Duarte-Rojo A, Guzmán-Valdivia-Gómez G, et al. Triggering receptor expressed on myeloid cells-1 expression on monocytes is associated with inflammation but not with infection in acute pancreatitis. Crit Care. 2009;13(3):R69. https://doi.org/10.1186/cc7876

24. Subramanian S, Pallati PK, Rai V, Sharma P, Agrawal DK, Nandipati KC. Increased expression of triggering receptor expressed on myeloid cells-1 in the population with obesity and insulin resistance. Obesity (Silver Spring). 2017;25(3):527-38. https://doi.org/10.1002/oby.21714

25. Lim SS, Norman RJ, Davies MJ, Moran LJ. The effect of obesity on polycystic ovary syndrome: a systematic review and meta-analysis. Obes Rev. 2013;14(2):95-109.

26. Guo Z, Jin F, Chen S, Hu P, Hao Y, Yu Q. Correlation between biochemical and clinical hyperandrogenism parameter in polycystic ovary syndrome in relation to age. BMC Endocr Disord. 2023;23(1):89. https://doi.org/10.1186/s12902-023-01346-x

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260287

# Artificial intelligence-driven triage in emergency departments: a comparative study of machine learning approaches

Ayşe Şule Akan[1]*, Sefa Kucuk[2], Şule Yakar[3], Mahmut Tutam[4], Hatice Esen[2], İbrahim Toker[3]

## SUMMARY

**OBJECTIVE:** In emergency departments, accurate and timely triage plays a critical role in prioritizing patient care. Traditional triage often suffers from subjectivity and inconsistency, especially under time pressure and overcrowded conditions. The aim of this study was to evaluate the predictive potential of multiple supervised machine learning models for three-class emergency department triage classification and compare their performance using routinely available triage variables.

**METHODS:** We conducted a comparative analysis of eight supervised machine learning models using a retrospective dataset comprising 510 patient records from the emergency department triage. The dataset included demographic features, vital signs, and chief complaints. A 10-fold cross-validation strategy was employed to evaluate performance in a three-class triage outcome scenario based on urgency levels using accuracy, precision, recall, and F1-score.

**RESULTS:** Among the evaluated models, logistic regression, random forest, support vector machine, and majority voting achieved the highest predictive performance, with overall accuracies exceeding 89%. The confusion matrix analysis revealed that low-urgency cases were consistently classified with the highest accuracy across all models, while misclassifications occurred primarily between high-urgency and medium-urgency cases.

**CONCLUSION:** The results show that machine learning models appear well-suited to support the emergency departments triage decision-making process. Integrating such models into emergency departments workflows can improve patient prioritization, streamline clinical operations, and ultimately contribute to better quality of care.

**KEYWORDS:** Emergency departments. Machine learning. Triage.

## INTRODUCTION

The recent global pandemic, coupled with an aging population, has led to a significant increase in demand for emergency medical assistance worldwide. In 2021, member countries of the Organization for Economic Cooperation and Development (OECD) with available data reported an average of 27 emergency department (ED) visits per 100 people annually[1]. Examples include Canada with 36.6 visits, England with 43.1, Australia with 34.2, and the United States of America with 43 visits per 100 people[2]. In Türkiye, the rate is significantly elevated, approximately 5.56 times the average of OECD countries[3]. When demand exceeds the capacity of EDs, several challenges in patient care can emerge. These include prolonged waiting times for medical evaluation, insufficient physical infrastructure, overburdened healthcare personnel, and delays in both treatment initiation and patient discharge[4]. In response to these challenges, emergency care in EDs typically begins with a triage process, which enables the prioritization of patients based on the severity and urgency of their medical needs.

Most widely adopted emergency department triage systems use five-level classification models. The most recognized are the Emergency Severity Index (ESI), Manchester Triage System (MTS), Canadian Triage and Acuity Scale (CTAS), and Australian Triage Scale (ATS), which form the basis of many global triage models. While most systems follow five levels, China uses a four-level system, the Chinese Emergency Triage Scale (CETS), whereas Türkiye uses a three-level, color-coded system called the Türkiye Triage System (TTS), established by a directive of the Turkish Ministry of Health in 2009[5]. TTS classifies patients into three levels: Red, Yellow, and Green, with Red and Yellow further divided into two subcategories[6]. Red represents the most critical cases: Subcategory 1 requires immediate intervention for life-threatening conditions (e.g., cardiac arrest, major trauma), while Subcategory 2 includes

[1]Erzurum City Hospital, Department of Emergency Medicine – Erzurum, Türkiye.
[2]Erzurum Technical University, College of Engineering and Architecture, Department of Electrical and Electronics Engineering – Erzurum, Türkiye.
[3]Kayseri City Hospital, Department of Emergency Medicine – Kayseri, Türkiye.
[4]Erzurum Technical University, College of Engineering and Architecture, Department of Industrial Engineering – Erzurum, Türkiye.
*Corresponding author: dr.aysesuleakan@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on March 01, 2026. Accepted on July 04, 2026.
Scientific Editor: José Maria Soares Júnior

high-risk cases needing treatment within 10 min (e.g., severe chest pain, acute neurological symptoms). Yellow covers serious but less urgent conditions: Subcategory 1 involves risk of limb loss or significant morbidity (e.g., severe hypertension, moderate bleeding), and Subcategory 2 includes prolonged yet non-critical issues such as minor trauma or gastrointestinal complaints. Green refers to stable patients with mild symptoms suitable for outpatient management.

Effective triage implementation ensures appropriate patient prioritization while optimizing ED resource allocation[7]. Reliable triage systems should be based on validated scoring methods[8]; however, variations in data collection, cultural context, and individual or environmental factors may introduce assessment errors[9]. Integrating artificial intelligence (AI) can mitigate these issues by enhancing the precision, consistency, and overall effectiveness of traditional triage systems[10]. Machine learning (ML) identifies actionable patterns within complex datasets through computational algorithms. For instance, a Gradient Boosted Decision Tree model implemented with the CatBoost library was developed to support ED triage officers in identifying potentially life-threatening mis-triage cases based on CETS[11]. This approach significantly reduced mis-triage among critically ill patients and outperformed logistic regression (LR).

Several ML models, Support Vector Machine (SVM), k-Nearest Neighbors (KNN), Decision Tree (DT), and Random Forest (RF), were evaluated for triage in the TTS[12], with RF achieving the best performance. Beyond outcome prediction, an XGBoost model was developed to recommend critical clinical interventions, improving timely decision-making[13]. An XGBoost-based model optimized with Hyperopt replicated clinicians' decisions, reducing over-triage and under-triage while improving efficiency[14]. Another study compared multiple ML algorithms for predicting discharge, hospitalization, ICU admission, and mortality, using oversampling and genetic algorithms for optimization; RF emerged as the top model[15]. Additionally, six ML models integrating structured and unstructured triage data via NLP were evaluated against clinicians' predictions, accurately forecasting ICU admission, mortality, and hospitalization, demonstrating their value in patient prioritization[16].

This study aims to compare the performance of several supervised ML algorithms for three-class ED triage prediction within the Turkish triage system using routinely available triage variables. By evaluating these models on a retrospective, single-center dataset, we seek to assess their predictive performance and identify methodological considerations for future externally validated triage decision-support studies.

# METHODS

## Dataset

This study used anonymized patient data collected from the ED of Kayseri City Hospital. Ethical approval for data collection was obtained from the Non-Interventional Clinical Research Ethics Committee. The data includes structured triage information for adult patients (aged 18 and above), including demographic characteristics (e.g., age, gender), vital signs, and clinical complaints. Only patients with complete triage data and approved urgency classifications were included in the analysis. The dataset comprises 510 patients equally represented in three triage categories: red, yellow, and green. The features were selected based on their routine availability during ED triage and relevance to urgency classification. The dataset includes structured information on patient demographics, vital signs, and chief complaints recorded at triage. Demographic variables comprise age and gender. Vital signs include systolic and diastolic blood pressure, heart rate, respiratory rate, body temperature, and oxygen saturation. Chief complaints represent the primary condition reported at admission, with a total of 58 distinct complaints recorded in the dataset.

## Data preparation

We applied several preprocessing steps to the dataset to ensure machine readability and consistency. Initially, categorical values such as gender and urgency labels were encoded into numerical representations using label encoding, allowing for direct processing by algorithms. To ensure comparability between variables, the six vital sign features were normalized between 0 and 1. The preprocessed data was exported in the attribute-relationship file format, which allows for seamless integration with the WEKA ML environment.

## Machine learning approaches

This study applied a supervised ML approach to train eight models on the ED triage dataset: BayesNet, DT, KNN, LR, Multilayer Perceptron (MLP), RF, SVM, and Majority Voting (MV). We employed model evaluation using WEKA. Given the relatively small size of the dataset, extensive hyperparameter search was avoided; instead, limited tuning was performed for the main tunable classifiers using WEKA's CVParameterSelection within the 10-fold cross-validation framework. BayesNet was kept with the standard WEKA configuration, and MV was reconstructed using optimized base classifiers. Table 1 summarizes the tuning strategy. Model performances were evaluated using accuracy, precision, recall, F1-score, specificity, and AUROC, and 95%CI for accuracy were calculated using the Wilson method.

## Implementation details

We performed data analysis and visualization in Python and classification experiments in WEKA due to its robust collection of ML algorithms and filters. Specifically, we applied the Partition Membership filter to transform the data into a propositionalized representation, as in[17]. This filter partitions the instance space using a partition generator based on DT models such as J48, RF, or REPTree and encodes membership values as new attributes. The resulting transformed instances are combined with the original class labels to form a sparse representation suitable for classification. This transformation allows traditional classifiers to be used on the data while preserving the essential structural information obtained from the partitioning process.

# RESULTS

In this section, we first explore the correlations between the features to understand their mutual relationships and then present the comparative performance of the ML models on the triage classification task.

Figure 1 presents the feature correlation matrix for the vital signs and TTS class. A positive correlation exists between diastolic and systolic blood pressure (r=0.69), aligning with physiological expectations. Heart rate shows a moderate positive correlation with respiratory rate (r=0.40), while oxygen saturation shows a moderate negative correlation with respiratory rate (r=-0.48). Correlations between individual vital signs and TTS class are generally weak to moderate, with the highest association observed for oxygen saturation (r=0.36), followed by respiratory rate (r=-0.20). The positive association between oxygen saturation and TTS suggests that higher saturation values are linked to less urgent triage categories, whereas the negative association of respiratory rate with TTS indicates higher respiratory rates in more urgent cases. Overall, no single vital sign strongly predicts triage class independently.

We evaluated the performance of several ML algorithms to classify patients into urgency levels. All models were evaluated using a 10-fold cross-validation strategy to obtain an internal estimate of predictive performance. In each iteration, nine folds were used for training, and the remaining fold was used for testing. This process was repeated until each fold served once as the test set, and the final performance scores were averaged across all folds.

The results in Table 2 show that LR and RF yield the highest overall classification performance in most metrics. LR achieves the highest accuracy (0.902; 95%CI 0.873–0.925), precision (0.902), and specificity (0.951), while RF closely follows and

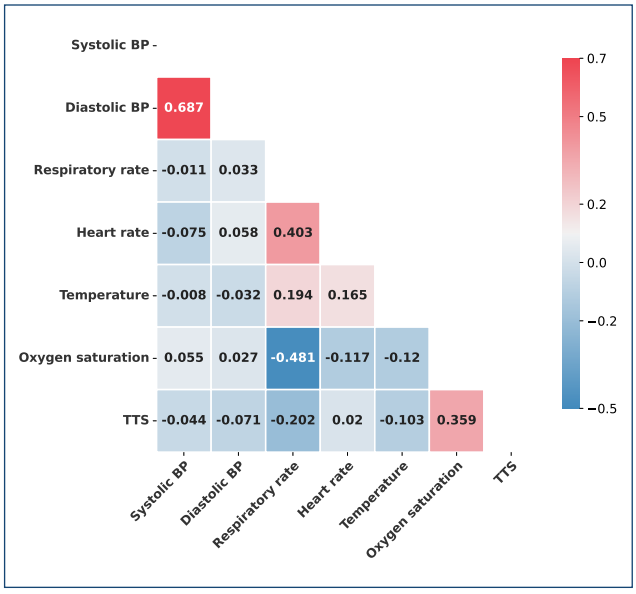



**Figure 1.** Correlation matrix of the features and Türkiye Triage System class.

**Table 1.** Hyperparameters used for the machine learning models.

| Models | Tuned parameter(s) | Search range | Selected value |
|---|---|---|---|
| BayesNet | Not tuned | Standard configuration | Default |
| DT | Confidence factor | 0.1–0.5 | 0.1 |
| KNN | Number of neighbors | 1–15 | 1 |
| LR | Ridge parameter | 1.0E-8–1.0E-4 | 1.0E-8 |
| MLP | Learning rate | 0.1–0.5 | 0.1 |
| RF | Number of trees | 50–300 | 50 |
| SVM | Complexity parameter | 0.1–10 | 2.575 |
| MV | Not independently tuned | Optimized base classifiers | Majority voting |

DT: Decision Tree; KNN: k-Nearest Neighbors; LR: logistic regression; MLP: Multilayer Perceptron; RF: Random Forest; SVM: Support Vector Machine; MV: Majority Voting.

achieves the highest AUROC. KNN, SVM, and MV also demonstrate competitive performance, with accuracy values close to 0.89. However, overlapping CI indicates small numerical differences among the best-performing models. In contrast, BayesNet underperforms across all metrics, with an accuracy of only 0.578 and an F1-score of 0.539, suggesting limited discriminative capacity on this dataset.

Figure 2 presents the confusion matrices, offering a detailed view of class-wise performance and highlighting prediction trends across urgency levels. LR, RF, SVM, and the MV ensemble demonstrate the most consistent and accurate results, with high true positive rates across all three TTS classes. In contrast, BayesNet performs substantially worse, frequently misclassifying TTS-2 and TTS-3, indicating limited ability to capture class overlap. DT and KNN show moderate performance. DT achieves reasonable accuracy, particularly for TTS-3, but often misclassifies TTS-1 as TTS-2, while KNN occasionally confuses TTS-1 and TTS-2, suggesting insufficient separation in local feature neighborhoods. MLP maintains a good balance across classes, though minor confusion between TTS-1 and TTS-2 persists, likely due to overlapping feature distributions.

The dataset is class-balanced, with 170 instances per urgency level. Among all classifiers, TTS-3 is the most accurately detected class: LR, RF, SVM, and MV correctly classify 162 of 170 TTS-3 cases. In contrast, TTS-2 is the most challenging class, with BayesNet misclassifying nearly half of its instances, often as TTS-3. Performance on TTS-1 is generally stable, though DT and MLP occasionally confuse it with TTS-2.

**Table 2.** Classification metrics comparison.

| Methods | Accuracy (95%CI) | Precision | Recall | F1-score | Specificity | AUROC |
|---|---|---|---|---|---|---|
| BayesNet | 0.578 (0.535–0.621) | 0.621 | 0.578 | 0.539 | 0.789 | 0.786 |
| DT | 0.845 (0.811–0.874) | 0.845 | 0.845 | 0.845 | 0.923 | 0.909 |
| KNN | 0.894 (0.864–0.918) | 0.895 | 0.894 | 0.894 | 0.947 | 0.935 |
| LR | 0.902 (0.873–0.925) | 0.902 | 0.902 | 0.902 | 0.951 | 0.931 |
| MLP | 0.880 (0.849–0.906) | 0.880 | 0.880 | 0.880 | 0.940 | 0.935 |
| RF | 0.898 (0.869–0.921) | 0.898 | 0.898 | 0.897 | 0.949 | 0.945 |
| SVM | 0.890 (0.860–0.914) | 0.890 | 0.890 | 0.890 | 0.945 | 0.929 |
| MV | 0.892 (0.862–0.916) | 0.892 | 0.892 | 0.892 | 0.946 | 0.919 |

DT: Decision Tree; KNN: k-Nearest Neighbors; LR: logistic regression; MLP: Multilayer Perceptron; RF: Random Forest; SVM: Support Vector Machine; MV: Majority Voting; AUROC: area under the receiver operating characteristic curve.

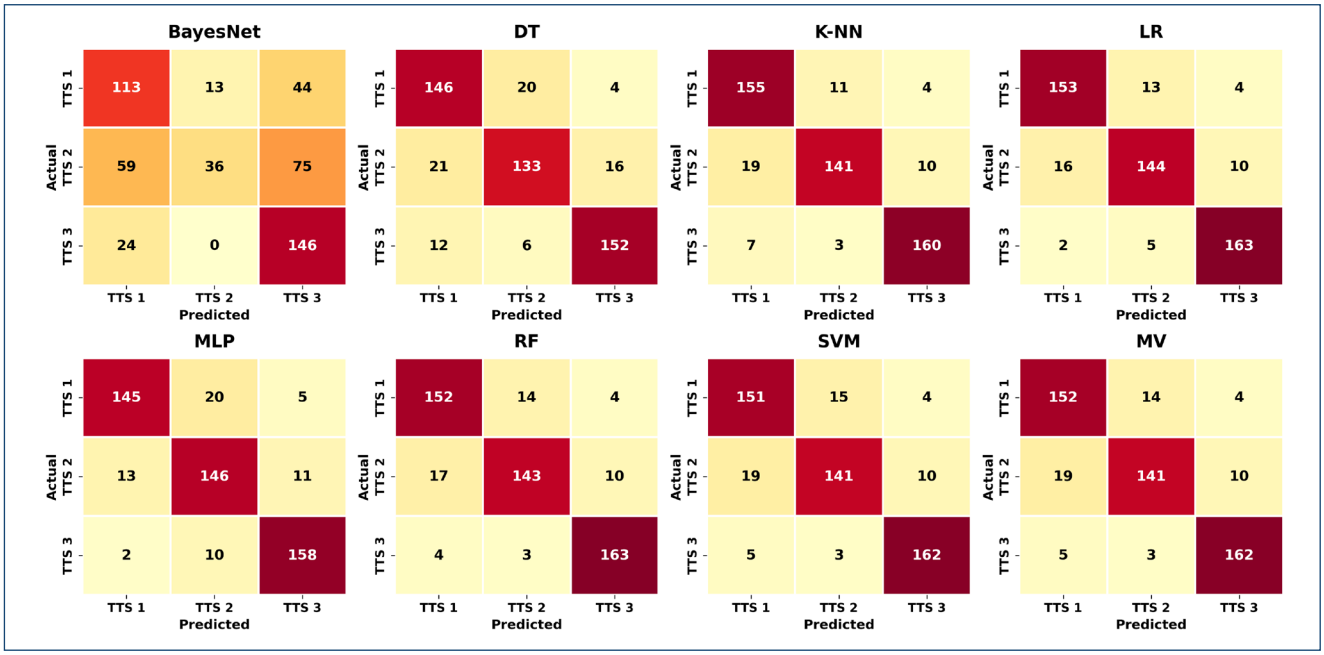



**Figure 2.** Confusion matrices obtained from a 10-fold cross-validation.

## DISCUSSION AND CONCLUSION

This study evaluated several ML algorithms for three-class ED triage prediction using a balanced single-center dataset. LR achieved the highest numerical accuracy, while RF showed the highest AUROC. However, overlapping Wilson CIs indicate that differences among the best-performing models should be interpreted with caution. Additionally, correlation analysis showed that routinely measured vital signs, particularly oxygen saturation, contained relevant information for TTS classification; however, no single variable could independently explain triage severity. These findings suggest that machine learning–based models can support triage decision-making processes but should be considered as complementary tools rather than replacements for clinical judgment.

Several limitations must be considered. The relatively small sample size, single-center retrospective design, and lack of external validation may restrict generalizability. While the study aimed to compare ML algorithms within a controlled and balanced dataset, further validation using larger, prospective, and multi-institutional cohorts is necessary to confirm robustness, fairness, and clinical applicability in real-world emergency care.

## AUTHORS' CONTRIBUTIONS

**AŞA:** Conceptualization, Writing – original draft. **SK:** Conceptualization, Methodology, Software, Visualization, Writing – review & editing. **ŞY:** Data curation. **MT:** Conceptualization, Methodology, Software, Writing – review & editing. **HE:** Software, Visualization. **İT:** Data curation.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. OECD. Health at a glance 2023: OECD indicators. Paris: OECD Publishing; 2023. https://doi.org/10.1787/7a7afb35-en
2. National Center for Health Statistics. National Hospital Ambulatory Medical Care Survey. 2022. Available from: https://www.cdc.gov/nchs/nhamcs/about/index.html
3. Birinci S, Ülgü MM, Gözükara MG. Critical insights based on the ministry of health's 6-year data analysis: an epidemiological study of patient visits trends of emergency departments in Türkiye. Haydarpasa Numune Train Res Hosp Med J. 2023;63(3):334-9. https://doi.org/10.14744/hnhj.2023.17048
4. Gottlieb M, Farcy DA, Moreno LA, Vilke GM, Guittard JA. Triage nurse-ordered testing in the emergency department setting: a review of the literature for the clinician. J Emerg Med. 2021;60(4):570-5. https://doi.org/10.1016/j.jemermed.2020.11.004
5. N. 27378 of the Republic of Turkey. Notification on the rules and principles of emergency services in in-patient health care facilities. 2009. Available from: https://www.resmigazete.gov.tr/eskiler/2009/10/20091016-16.htm
6. Aslanoğlu A, Ayyıldız M. Triage in the world and in Turkey and the triage knowledge levels of emergency ward nurses. Sağlık Bilim Derg. 2023;32(2) :316-22. https://doi.org/10.34108/eujhs.1159047
7. Andersson J, Nordgren L, Cheng I, Nilsson U, Kurland L. Long emergency department length of stay: a concept analysis. Int Emerg Nurs. 2020;53:100930. https://doi.org/10.1016/j.ienj.2020.100930
8. Cairós-Ventura LM, Mercedes Novo-Muñoz M, Rodríguez-Gómez JÁ, Ortega-Benítez ÁM, Ortega-Barreda EM, Aguirre-Jaime A. Validity and reliability of the emergency severity index in a Spanish hospital. Int J Environ Res Public Health. 2019;16(22):4567. https://doi.org/10.3390/ijerph16224567
9. Souza CC, Toledo AD, Tadeu LF, Chianca TC. Risk classification in an emergency room: agreement level between a Brazilian institutional and the Manchester Protocol. Rev Lat Am Enfermagem. 2011;19(1):26-33. https://doi.org/10.1590/s0104-11692011000100005
10. Stewart J, Sprivulis P, Dwivedi G. Artificial intelligence and machine learning in emergency medicine. Emerg Med Australas. 2018;30(6):870-4. https://doi.org/10.1111/1742-6723.13145
11. Liu Y, Gao J, Liu J, Walline JH, Liu X, Zhang T, et al. Development and validation of a practical machine-learning triage algorithm for the detection of patients in need of critical care in the emergency department. Sci Rep. 2021;11(1):24044. https://doi.org/10.1038/s41598-021-03104-2
12. Bora Ş, Kantarci A, Erdogan A, Beynek B, Kheibari B, Evren V, et al. Machine learning for E-triage. Int J Multidiscip Stud Innov Technol. 2022;6(1) :86. https://doi.org/10.36287/ijmsit.6.1.86
13. Chang H, Yu JY, Yoon S, Kim T, Cha WC. Machine learning-based suggestion for critical interventions in the management of potentially severe conditioned patients in emergency department triage. Sci Rep. 2022;12(1):10537. https://doi.org/10.1038/s41598-022-14422-4
14. Gao Z, Qi X, Zhang X, Gao X, He X, Guo S, et al. Developing and validating an emergency triage model using machine learning algorithms with medical big data. Risk Manag Healthc Policy. 2022;15:1545-51. https://doi.org/10.2147/RMHP.S355176
15. Elhaj H, Achour N, Tania MH, Aciksari K. A comparative study of supervised machine learning approaches to predict patient triage outcomes in hospital emergency departments. Array. 2023;17:100281. https://doi.org/10.1016/j.array.2023.100281
16. Chang YH, Lin YC, Huang FW, Chen DM, Chung YT, Chen WK, et al. Using machine learning and natural language processing in triage for prediction of clinical disposition in the emergency department. BMC Emerg Med. 2024;24(1):237. https://doi.org/10.1186/s12873-024-01152-1
17. Daşdemir Y, Özakar R. Affective states classification performance of audio-visual stimuli from EEG signals with multiple-instance learning. Turkish J Electr Eng Comput Sci. 2022;30(7):2707-24.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260298

# Association between triglyceride–glucose and fibrosis-4 indices and nephrolithiasis: a cross-sectional case–control study

Sezin Doğan Çakır[1]*, Mustafa Erkoç[2], Mine Adaş[1]

## SUMMARY

**OBJECTIVE:** Fibrosis-4 (a liver fibrosis index) and triglyceride–glucose (an insulin resistance surrogate) indices may be markers associated with kidney stone risk. The aim of this study was to evaluate the association between triglyceride–glucose and fibrosis-4 indices and the presence of nephrolithiasis in adults.

**METHODS:** This retrospective cross-sectional case–control study included 918 adults (785 with nephrolithiasis, 133 controls) evaluated between January 2016 and December 2023. Controls were consecutive patients from the same urology outpatient population who underwent abdominal imaging for non-stone-related urological indications, with no radiological evidence of nephrolithiasis. Triglyceride–glucose and fibrosis-4 indices were calculated using standard formulas. The triglyceride–glucose index was standardized (per 1 SD=0.23 units) prior to multivariable logistic regression including age, sex, and fibrosis-4.

**RESULTS:** Triglyceride–glucose was significantly higher in stone formers (4.87±0.21 vs. 4.64±0.20; p<0.001). In multivariable analysis, triglyceride–glucose independently predicted nephrolithiasis (adjusted OR per 1 SD=3.52; 95%CI 2.48–5.00; p<0.001); male sex was also independently associated. Fibrosis-4 was not independently associated (p=0.809). Triglyceride–glucose area under the curve was 0.779 (95%CI 0.729–0.827), consistent across both sexes (interaction p=0.135).

**CONCLUSION:** Each 1 SD increase in triglyceride–glucose was associated with a 3.5-fold increase in odds of nephrolithiasis, independent of age, sex, and fibrosis-4. Fibrosis-4 showed no independent utility. Prospective studies are needed to confirm these findings.

**KEYWORDS:** Nephrolithiasis. Insulin resistance. Triglycerides. Blood glucose. Biomarkers.

## INTRODUCTION

Nephrolithiasis affects 10–15% of adults worldwide and poses a growing clinical and economic burden[1]. Increasing evidence supports the concept of kidney stone disease as a systemic metabolic condition rather than an isolated urological disorder[2,3]. Metabolic syndrome and its components, particularly insulin resistance and diabetes, have been consistently associated with nephrolithiasis in population-based studies and meta-analyses, reinforcing a systemic metabolic basis for stone formation[4,5].

Metabolic dysfunction-associated steatotic liver disease (MASLD) has similarly emerged as a major global health concern, currently affecting more than one-third of adults, with prevalence projected to exceed 55% by 2040[6]. Meta-analyses have confirmed an association between MASLD and nephrolithiasis, suggesting a shared metabolic and inflammatory pathophysiology linking hepatic and renal manifestations of insulin resistance[7,8].

The triglyceride–glucose (TyG) index, derived from fasting triglyceride and glucose levels, is a validated surrogate marker of insulin resistance and has been associated with nephrolithiasis[9,10]. The fibrosis-4 (FIB-4) index, calculated using age, aminotransferases, and platelet count, is a first-line noninvasive tool for assessing liver fibrosis in MASLD and reflects cumulative metabolic burden[11]. Data evaluating FIB-4 in stone formers remain limited. This study aimed to assess the association between TyG and FIB-4 indices and nephrolithiasis in adults.

## METHODS

### Study design and population

This retrospective cross-sectional case–control study was conducted at Prof. Dr. Cemil Taşcıoğlu City Hospital. Adults (≥18 years) who underwent abdominal imaging in the urology department between

[1]University of Health Sciences, Prof. Dr. Cemil Taşcıoğlu City Hospital, Department of Endocrinology – Istanbul, Turkey.
[2]University of Health Sciences, Prof. Dr. Cemil Tascıoğlu City Hospital, Department of Urology – Istanbul, Turkey.
*Corresponding author: sezin.dogan@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 11, 2026. Accepted on July 24, 2026.
Scientific Editor: José Maria Soares Júnior

January 2016 and December 2023 were screened. All consecutive eligible patients meeting the inclusion criteria were included. Patients with radiologically confirmed nephrolithiasis constituted the stone group. Controls were selected from the same source population and consisted of consecutive patients who underwent abdominal ultrasonography and/or computed tomography for non-stone-related urological indications, including microscopic hematuria, non-specific flank pain, lower urinary tract symptoms, simple renal cysts, varicocele, or hydrocele, with no imaging evidence of nephrolithiasis.

Participants with complete metabolic laboratory and imaging data were included. Exclusion criteria applied to both groups were chronic viral or autoimmune liver disease, diabetes mellitus, pregnancy, malignancy, chronic kidney disease, active infection (including urinary tract infection, pyelonephritis, and prostatitis), clinically significant urinary obstruction, and incomplete records. Given the retrospective nature of the study and the use of de-identified patient data, the ethics committee granted a waiver of individual informed consent, in accordance with institutional policy and the Declaration of Helsinki. The study was approved by the local ethics committee (Approval date: 20.11.2023; Approval number: 273).

### Data collection

Demographic characteristics, comorbidities, and fasting laboratory parameters (triglycerides, glucose, aspartate aminotransferase (AST), alanine aminotransferase (ALT), and platelet count) were obtained from electronic medical records. Abdominal imaging included non-contrast computed tomography and/or ultrasonography. The presence of hepatic steatosis was recorded based on standard radiologic criteria.

### Index calculations

FIB-4=(Age×AST) / [Platelet count×√ALT][11]. TyG=ln [fasting triglycerides (mg/dL)×fasting glucose (mg/dL)/2][9].

### Statistical analysis

Continuous variables were expressed as mean±SD, and categorical variables as frequencies and percentages. Normality was assessed using the Kolmogorov-Smirnov test. Triglycerides and FIB-4 were non-normally distributed and were compared using the Mann-Whitney U test; other variables were compared using independent samples t-tests. Categorical variables were compared using chi-square tests. Multivariable logistic regression was performed using the enter method with age, sex, TyG index, and FIB-4 index entered simultaneously. The TyG index was standardized (z-score: [TyG-mean]/SD) prior to analysis to yield odds ratios per 1 SD. Sex-stratified analyses and a sex×TyG interaction term were tested. Receiver operating characteristic (ROC) curve analysis evaluated discriminative performance of TyG; area under the curve (AUC) was calculated with 95%CIs (bootstrap, 2000 iterations). Statistical significance was $p<0.05$ (two-tailed). SPSS version 25.0 (IBM Corp., Armonk, NY, USA) was used.

## RESULTS

### Baseline characteristics

A total of 918 individuals were included (785 stone formers, 133 controls). Age did not differ significantly between groups (45.77±11.55 vs. 44.29±9.88 years; p=0.161). Male sex was more prevalent among stone formers (69.6 vs. 37.6%; p<0.001). Comorbidities were comparable (Table 1).

### Laboratory parameters and metabolic indices

The TyG index was significantly higher in the nephrolithiasis group (4.87±0.21 vs. 4.64±0.20; p<0.001), driven by markedly higher fasting triglyceride levels (189.16±70.11 vs. 127.31±47.02 mg/dL; p<0.001). Fasting glucose did not differ significantly (p=0.101). The FIB-4 index was modestly elevated in stone formers (0.83±0.38 vs. 0.77±0.27; p=0.040); however, both values are well within the low-risk range for advanced hepatic fibrosis (<1.45), and this difference did not survive Bonferroni correction (adjusted threshold ≈0.006 for eight comparisons). Aminotransferases and platelet counts were similar between groups (Table 2).

**Table 1.** Baseline demographic and clinical characteristics.

| Variable | Nephrolithiasis (n=785) | Controls (n=133) | p-value |
|---|---|---|---|
| Age (years) | 45.77±11.55 | 44.29±9.88 | 0.161 |
| Male sex, n (%) | 546 (69.6%) | 50 (37.6%) | <0.001 |
| Hypertension, n (%) | 13.1% | 13.2% | 0.946 |
| Cardiovascular disease, n (%) | 1.4% | 0.0% | 0.553 |
| Inflammatory/rheumatologic disease, n (%) | 0.4% | 0.0% | 0.763 |

Continuous variables: mean±SD; categorical: n (%).

## Multivariable analysis

The TyG index was standardized prior to analysis (mean=4.826, SD=0.227; the observed group difference of 0.226 units corresponds to 1.00 SD). In multivariable logistic regression, TyG was the only independent predictor of nephrolithiasis (adjusted OR per 1 SD=3.52; 95%CI 2.48–5.00; p<0.001). Male sex was also independently associated (OR 11.42; 95%CI 5.67–23.01; p<0.001). Age (OR 0.999; p=0.963) and FIB-4 (OR 0.885; p=0.809) were not independently associated. The model achieved 83.3% overall accuracy (Nagelkerke pseudo-$R^2$=0.511; Table 3).

## Receiver operating characteristic analysis

ROC analysis (Figure 1) demonstrated moderate discriminatory ability of the TyG index (AUC=0.779; 95%CI 0.729–0.827; p<0.001). The FIB-4 index showed no meaningful discriminatory ability (AUC=0.545; 95%CI 0.479–0.606; p=0.167). A TyG cutoff of 4.69 yielded 79.5% sensitivity and 61.2% specificity, with a positive predictive value of 89.8% and a negative predictive value of 40.9%. The high positive predictive value reflects the high prevalence of nephrolithiasis in this clinical sample (~85.5%) and may not generalize to lower-prevalence populations, whereas the low negative predictive value indicates limited rule-out performance.

## Sex-stratified analysis

Sex-stratified analysis demonstrated that the association between TyG and nephrolithiasis remained significant in both males (OR 5.14; 95%CI 2.27–11.66; p<0.001; AUC=0.880) and females (OR 3.09; 95%CI 2.09–4.56; p<0.001; AUC=0.761).

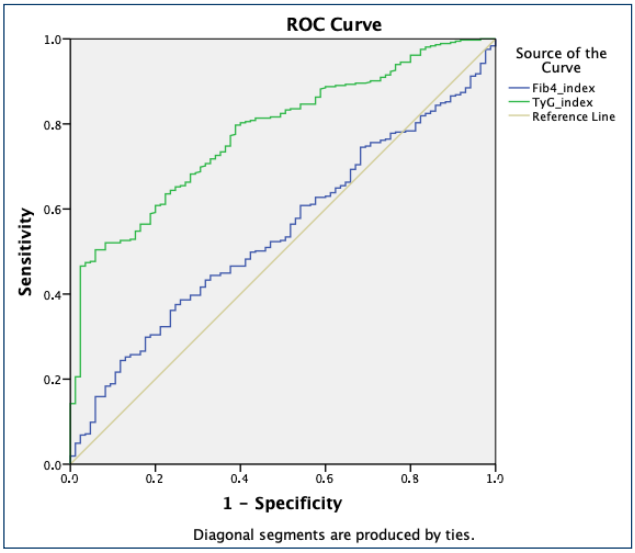



**Figure 1.** Receiver operating characteristic curve of the triglyceride–glucose and fibrosis-4 index for predicting nephrolithiasis. The area under the curve was 0.779 for triglyceride–glucose (p<0.001), while the area under the curve was 0.545 for fibrosis-4 (p=0.375).

**Table 2.** Laboratory parameters and metabolic indices by group.

| Variable | Nephrolithiasis (n=785) | Controls (n=133) | p-value |
|---|---|---|---|
| Fasting glucose (mg/dL) | 96.2±39.4 | 90.6±7.7 | 0.101 |
| Triglycerides (mg/dL) | 189.16±70.11 | 127.31±47.02 | <0.001 |
| TyG index | 4.87±0.21 | 4.64±0.20 | <0.001 |
| AST (U/L) | 19.8±6.3 | 19.2±4.8 | 0.311 |
| ALT (U/L) | 20.7±9.3 | 19.6±6.5 | 0.118 |
| Platelet count (×$10^9$/L) | 261.11±54.21 | 258.77±45.08 | 0.593 |
| FIB-4 index | 0.83±0.38 | 0.77±0.27 | 0.040† |

†p=0.040 for FIB-4 does not survive Bonferroni correction (adjusted threshold ≈0.006). Triglycerides and FIB-4: Mann-Whitney U test; others: independent samples t-test. TyG: triglyceride–glucose; FIB-4: fibrosis-4; AST: aspartate aminotransferase; ALT: alanine aminotransferase.

**Table 3.** Multivariable logistic regression for nephrolithiasis.

| Variable | Adjusted OR | 95%CI | p-value |
|---|---|---|---|
| TyG index (per 1 SD) | 3.52 | 2.48–5.00 | <0.001 |
| Sex (male vs. female) | 11.42 | 5.67–23.01 | <0.001 |
| Age (per year) | 0.999 | 0.967–1.033 | 0.963 |
| FIB-4 index | 0.885 | 0.328–2.383 | 0.809 |

Enter method; TyG standardized (per 1 SD=0.23 units). Overall accuracy: 83.3%; Nagelkerke pseudo-$R^2$=0.511. OR: odds ratio; CI: confidence interval; TyG: triglyceride–glucose; SD: standard deviation; FIB-4: fibrosis-4.

The sex×TyG interaction term was not statistically significant (OR 1.95; 95%CI 0.81–4.68; p=0.135), indicating no significant effect modification by sex.

## DISCUSSION

This cross-sectional study demonstrates that both TyG and FIB-4 indices are elevated in individuals with nephrolithiasis, supporting the concept of kidney stone disease as a manifestation of systemic metabolic dysfunction. The TyG index showed a stronger cross-sectional association and superior discriminative performance.

The association between elevated TyG and nephrolithiasis is consistent with previous large population-based studies[10,12,13]. Each 1 SD increase in TyG (0.23 units, approximately the observed group difference) was independently associated with a 3.5-fold increase in the odds of nephrolithiasis. This standardized estimate is more clinically interpretable than raw-unit odds ratios and is consistent with the graded risk relationships reported in NHANES-based and Chinese population-based case–control studies[10,12]. Our TyG cutoff of 4.69 aligns with the range reported in prior studies (approximately 4.5–5.0).

The TyG–nephrolithiasis association was consistent across both sexes (male OR 5.14, AUC=0.880; female OR 3.09, AUC=0.761) with no significant sex interaction (p=0.135), in keeping with NHANES-based findings[10]. The somewhat higher OR in males may reflect the small male control group (n=14) and should be interpreted cautiously. The marked male predominance among stone formers (69.6 vs. 37.6%) is consistent with established epidemiology[2].

Mechanistically, insulin resistance provides a plausible biological link. Impaired renal ammoniagenesis and reduced urinary citrate excretion lead to lower urinary pH, promoting calcium oxalate and uric acid crystallization[14,15]. Insulin resistance is also associated with tubular inflammation, oxidative stress, and altered electrolyte handling, which may further facilitate lithogenesis[16].

The modest elevation of FIB-4 among stone formers aligns with emerging evidence linking MASLD and nephrolithiasis[17,18]. Meta-analyses have confirmed a significant association between MASLD and kidney stone disease, while imaging-based investigations have demonstrated increased hepatic steatosis prevalence among nephrolithiasis patients, supporting shared metabolic pathways[19]. However, the FIB-4 difference (0.83 vs. 0.77) is clinically negligible—both values are within the low-risk range (<1.45)—the difference did not survive Bonferroni correction, and FIB-4 was not independently associated (OR 0.885; p=0.809). Its AUC of 0.545 indicates no meaningful discriminatory ability beyond chance, suggesting that hepatic fibrotic burden is not an independent driver of nephrolithiasis in this population.

Notably, aminotransferases did not differ significantly between groups, highlighting the value of composite indices such as FIB-4 in detecting subclinical metabolic alterations, even when individual components are unremarkable. This study provides real-world clinical data regarding the association of insulin resistance-based and fibrosis-based indices with nephrolithiasis.

Several limitations should be acknowledged. The cross-sectional design precludes causal inference. The unequal case-to-control ratio (~6:1) reflects the clinical setting of a tertiary referral urology center and may inflate sex-specific OR estimates and affect PPV/NPV generalizability; residual selection bias cannot be excluded. Dietary factors, urinary biochemistry, stone composition, and stone location were not systematically available. BMI was available only in the stone group (n=209; mean 28.69±4.21 kg/m$^2$) and absent in controls. Although hepatic steatosis was recorded descriptively from available imaging reports, these data were incomplete across the study population and therefore were not incorporated into multivariable analyses. Medication use and family history of stones were not recorded. Advanced imaging for liver fibrosis validation also was not performed.

Future prospective studies should evaluate whether TyG predicts incident stone formation and recurrence, and whether insulin resistance-targeted interventions reduce stone risk.

## CONCLUSION

In this cross-sectional case–control study, each 1 SD increase in TyG index was independently associated with a 3.5-fold increase in the odds of nephrolithiasis (adjusted OR 3.52; 95%CI 2.48–5.00; p<0.001), with moderate discriminatory ability (AUC=0.779). This association was consistent across both sexes with no significant sex interaction. FIB-4 did not demonstrate independent association or meaningful discriminatory ability. Prospective longitudinal studies are needed to determine whether TyG predicts incident nephrolithiasis and to clarify its potential clinical applicability.

## AUTHORS' CONTRIBUTIONS

**SDÇ:** Data curation, Formal Analysis, Writing – original draft. **ME:** Data curation, Formal Analysis. **MA:** Data curation, Formal Analysis.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Carvalho M, Matos ACC, Santos DRD, Barreto DV, Barreto FC, Rodrigues FG, et al. Brazilian Guidelines on evaluation and clinical management of nephrolithiasis: Brazilian Society of Nephrology. J Bras Nefrol. 2025;47(2):e20240189. https://doi.org/10.1590/2175-8239-JBN-2024-0189en
2. Bhojani N, Bjazevic J, Wallace B, Lee L, Kaler KS, Dion M, et al. UPDATE - Canadian Urological Association guideline: evaluation and medical management of kidney stones. Can Urol Assoc J. 2022;16(6):175-88. https://doi.org/10.5489/cuaj.7872
3. Sakhaee K, Maalouf NM, Sinnott B. Clinical review. Kidney stones 2012: pathogenesis, diagnosis, and management. J Clin Endocrinol Metab. 2012;97(6):1847-60. https://doi.org/10.1210/jc.2011-3492
4. Chen L, Zhang J, Shen K, Zhu Y, Zhang J, Pan J, et al. Kidney stones are associated with metabolic syndrome in a health screening population: a cross-sectional study. Transl Androl Urol. 2023;12(6):967-76. https://doi.org/10.21037/tau-23-51
5. Dassanayake SN, Lafont T, Somani BK. Association and risk of metabolic syndrome and kidney stone disease: outcomes from a systematic review and meta-analysis. Curr Opin Urol. 2025;35(4):377-84. https://doi.org/10.1097/MOU.0000000000001227
6. Younossi ZM, Kalligeros M, Henry L. Epidemiology of metabolic dysfunction-associated steatotic liver disease. Clin Mol Hepatol. 2025;31(Suppl):S32-50. https://doi.org/10.3350/cmh.2024.0431
7. Qin S, Wang S, Wang X, Wang J. Non-alcoholic fatty liver disease and the risk of urolithiasis: a systematic review and meta-analysis. Medicine (Baltimore). 2018;97(35):e12092. https://doi.org/10.1097/MD.0000000000012092
8. Mantovani A, Morandin R, Fiorio V, Lando MG, Petta S, Ferraro PM, et al. Association between metabolic dysfunction-associated steatotic liver disease and risk of urolithiasis: an updated systematic review and meta-analysis. Intern Emerg Med. 2024;19(6):1745-55. https://doi.org/10.1007/s11739-024-03705-5
9. Simental-Mendía LE, Rodríguez-Morán M, Guerrero-Romero F. The product of fasting glucose and triglycerides as surrogate for identifying insulin resistance in apparently healthy subjects. Metab Syndr Relat Disord. 2008;6(4):299-304. https://doi.org/10.1089/met.2008.0034
10. Qin Z, Zhao J, Geng J, Chang K, Liao R, Su B. Higher triglyceride-glucose index is associated with increased likelihood of kidney stones. Front Endocrinol (Lausanne). 2021;12:774567. https://doi.org/10.3389/fendo.2021.774567
11. European Association for the Study of the Liver (EASL); European Association for the Study of Diabetes (EASD); European Association for the Study of Obesity (EASO). EASL-EASD-EASO Clinical Practice Guidelines on the management of metabolic dysfunction-associated steatotic liver disease (MASLD). J Hepatol. 2024;81(3):492-542. https://doi.org/10.1016/j.jhep.2024.04.031
12. Wang D, Zhang D, Zhang L, Shi F, Zhu Y. Association between triglyceride-glucose index and risk of kidney stone: a Chinese population-based case-control study. BMJ Open. 2024;14(11):e086641. https://doi.org/10.1136/bmjopen-2024-086641
13. Yu H, Wu J. Associations of triglyceride glucose-related parameters with kidney stones: a cross-sectional study from NHANES 2007-2020. Transl Androl Urol. 2025;14(2):379-88. https://doi.org/10.21037/tau-24-516
14. Cunneely OP, Roberts A, Fargue S, Knight J, Assimos DG, Wood KD. Metabolic dysfunction associated steatotic liver and kidney stones: what is going on? Curr Opin Nephrol Hypertens. 2025;34(3):247-53. https://doi.org/10.1097/MNH.0000000000001062
15. Shastri S, Patel J, Sambandam KK, Lederer ED. Kidney stone pathophysiology, evaluation and management: core curriculum 2023. Am J Kidney Dis. 2023;82(5):617-34. https://doi.org/10.1053/j.ajkd.2023.03.017
16. Xu Z, Yao X, Duan C, Liu H, Xu H. Metabolic changes in kidney stone disease. Front Immunol. 2023;14:1142207. https://doi.org/10.3389/fimmu.2023.1142207
17. Qin S, Wang J, Zhou C, Zhang Y, Xu Y, Wang X, et al. The association between a non-invasive hepatic fibrosis score and urolithiasis among non-alcoholic fatty liver disease (NAFLD) patients in China: a cross-sectional study. BMJ Open. 2019;9(8):e027702. https://doi.org/10.1136/bmjopen-2018-027702
18. Zhang F, Li W. Association between the Fatty Liver Index, metabolic dysfunction-associated steatotic liver disease, and the risk of kidney stones. Kidney Blood Press Res. 2025;50(1):115-30. https://doi.org/10.1159/000543404
19. Nguyen DD, Bouhadana D, Wong P, Andonian S. What is the prevalence of hepatic steatosis on ultrasonography in patients followed for nephrolithiasis? Can Urol Assoc J. 2022;16(12):442-6. https://doi.org/10.5489/cuaj.7767

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260317

# Postural imbalance in patients with rheumatoid arthritis: a cross-sectional study

Ana Luisa Trevisan[1], Bruna Furukawa[1], Clara Mayumi Mikami[1], Naira Alessandra Hurmus[1], Thelma Larocca Skare[1], Renato Nisihara[1]*, Barbara Stadler Kahlow[1]

## SUMMARY

**OBJECTIVE:** The aim of this study was to evaluate postural balance in patients with rheumatoid arthritis and identify factors associated with balance impairment.
**METHODS:** This cross-sectional case–control study included 101 patients with rheumatoid arthritis and 94 healthy controls recruited at a single tertiary referral center between March 2024 and April 2025. Epidemiological and clinical data, including disease activity, current treatments, and the number of falls during the previous year, were obtained through medical record review and direct patient interviews. Balance was assessed using the Berg Balance Scale, and handgrip strength was measured with a Jamar® dynamometer.
**RESULTS:** Patients with rheumatoid arthritis had significantly lower Berg Balance Scale scores (p<0.0001) and a markedly higher likelihood of experiencing falls than controls (OR 6.2; 95%CI 1.48–28.2). In univariate analysis, age (p<0.0001), handgrip strength (p=0.002), hypertension (p=0.02), tobacco exposure (p=0.0008), antimalarial use (p=0.02), and methotrexate use (p=0.007) were significantly associated with Berg Balance Scale scores. However, after multivariable adjustment, only antimalarial use (β=-2.75;95%CI -5.12 to -0.38; p=0.02), handgrip strength (β=0.09; 95%CI 0.01–0.17; p=0.01), and tobacco exposure (β=-1.90; 95%CI -3.20 to -0.62; p=0.004) remained independently associated with balance performance.
**CONCLUSION:** Postural imbalance appears to be an important and potentially modifiable feature of rheumatoid arthritis. Routine balance assessment, together with interventions targeting muscle strength and smoking cessation, may help reduce fall risk and improve functional outcomes. Further studies are needed to clarify the relationship between antimalarial use and balance impairment.
**KEYWORDS:** Rheumatoid arthritis. Postural balance. Accidental falls.

## INTRODUCTION

Postural imbalance is a clinically relevant but still underexplored manifestation of rheumatoid arthritis (RA). Balance, a sophisticated neuromuscular function, refers to the body's ability to control its center of gravity with minimal postural sway and maximum stability within the base of support[1]. Impaired postural control leads to increased postural sway, which, even in healthy individuals, is associated with a higher risk of falls[2]. In patients with RA, balance may be impaired by multiple mechanisms, including reduced joint proprioception secondary to foot deformities and lower-limb arthritis, muscle weakness, limited joint mobility, and, in some cases, central nervous system involvement[1]. Consequently, individuals with RA have a higher risk of falls than healthy individuals[1]. This elevated fall risk is further exacerbated by factors such as the use of antidepressants, impaired vision, and antihypertensive medications[1]. The concern surrounding falls in RA patients is amplified by their frequently compromised bone health and an increased propensity for fractures, largely attributable to reduced bone mass and the high prevalence of osteoporosis within this population[3]. In RA, bone loss is associated with the underlying inflammatory process, characterized by elevated levels of TNF-alpha, IL-6, and IL-1, which promote osteoclast activation and bone resorption. Moreover, the widespread use of glucocorticoids in RA management, although effective for controlling inflammation, is another significant risk factor for the development of secondary osteoporosis. Despite major advances in RA treatment, particularly the widespread use of disease-modifying antirheumatic drugs (DMARDs) and efforts to minimize glucocorticoid exposure, fracture risk remains substantial. A prospective observational study conducted in Japan demonstrated an increase in both vertebral and non-vertebral fractures among patients with RA between 2011 and 2023[4].

A meta-analysis including 34 studies and 24,123 participants reported a pooled prevalence of 34% (95%CI 29–38%) for at least one fall and 16% (95%CI 12–20%) for recurrent

[1]Mackenzie Evangelical School of Medicine of Paraná – Curitiba (PR), Brazil.
*Corresponding author: renatonisihara@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on July 18, 2026. Accepted on July 20, 2026.
Scientific Editor: Roseli Nomura

falls among adults with RA[5]. In another study involving 4,996 individuals with RA, 10.1% reported experiencing at least one fall, 2.2% reported multiple falls, and 19.2% expressed a fear of falling within the previous 6 months[6]. Notably, those who reported falls were significantly more likely to have sustained fractures during the same period compared to non-fallers (14.7 versus 1.1%)[1].

Given this background and the scarcity of data on postural balance among Brazilian patients with RA, we evaluated postural balance and fall risk in a cohort of individuals with RA. Additionally, we aimed to examine potential associations with age, disease duration, disease activity, muscle strength, and comorbid conditions. Generating local data is particularly important because genetic background, environmental factors, and disparities in access to healthcare may influence disease severity, functional outcomes, and fall risk in patients with RA.

## METHODS

This is a cross-sectional case–control study with a convenience sample of RA patients from a single Rheumatology Center that cares for patients from the Public Health Service in Brazil. Patients were recruited from March 2024 to April 2025. The project was approved by the Institutional Committee of Ethics in Research under protocol number 6.759.253, and included patients were invited to participate according to appointment order. To be included, the patients should fulfill at least 6 points in the American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) Classification criteria for RA[7], be over 18 years of age, and provide written informed consent. Individuals with neurological disorders, lower-limb orthopedic conditions, or amputations were excluded. Participants were consecutively recruited according to the order of their scheduled outpatient consultations, provided they met the predefined inclusion and exclusion criteria and expressed willingness to participate in the study. Companions accompanying the patients were invited to serve as controls and were matched to the study group by sex and age.

### Data collection included

Epidemiological and clinical data: age, sex, self-declared ethnic background, tobacco use, disease duration, presence of rheumatoid factor, current medications, comorbidities, and the number of falls in the previous year. Information was obtained through medical chart review and direct patient interviews.

Disease activity assessment: Disease activity was evaluated using the DAS28-ESR (Disease Activity Score using 28 joints and erythrocyte sedimentation rate), CDAI (Clinical Disease Activity Index), and SDAI (Simplified Disease Activity Index)[8]. The DAS28-ESR incorporates the number of tender and swollen joints (out of 28), the ESR value, and the patient's general health status or global disease activity assessed via a 100 mm visual analog scale (VAS). The CDAI is based on the 28-joint tender and swollen counts, the patient's global assessment of disease activity (0–10 scale), and the evaluator's global assessment (0–10 scale). The SDAI includes the same parameters as the CDAI, with the addition of the C-reactive protein (CRP) level[8].

Functional balance evaluation: The Berg Balance Scale (BBS) was a validated instrument used to assess patients' static and dynamic balance across 14 tasks. Each item is scored on a scale from 0 to 4, yielding a maximum total score of 56 points[9]. Scores below 45 suggest increased risk of falling, with scores under 40 being associated with an almost certain risk of falls[10,11].

Handgrip strength measurement: Handgrip strength was assessed using a Jamar® dynamometer, with participants seated and elbows flexed at 90 degrees. In total, three measurements were taken, and the highest value was used for analysis. Measurements were performed on the dominant hand.

Statistical analysis: Data were organized in frequency and contingency tables. The Shapiro-Wilk test was applied to determine the distribution of continuous variables. Comparisons of BBS scores across clinical, epidemiological, and therapeutic subgroups were performed using the Mann-Whitney U test. Correlations between BBS and numerical variables were evaluated using Spearman's rank correlation coefficient. A p-value of $<0.05$ was considered statistically significant. Variables that demonstrated an association with BBS at a significance level of $p<0.1$ were subsequently entered into a multivariate analysis to assess their independence. All statistical computations were carried out using MedCalc® version 10.0.

## RESULTS

A total of 101 patients with RA and 94 healthy controls were included in the study. None were excluded. Table 1 summarizes the epidemiological characteristics of patients and controls, the clinical characteristics of the RA group, and the results of the BBS and handgrip strength assessments. Compared with controls, patients with RA had significantly lower BBS scores, reported more falls, and exhibited lower handgrip strength.

Table 2 presents the correlations between BBS scores and continuous variables. Age showed a moderate negative correlation with BBS scores, whereas handgrip strength demonstrated a modest positive correlation.

Table 3 presents the distribution of BBS scores according to categorical variables. Lower BBS scores were significantly

associated with tobacco exposure, antimalarial use, methotrexate use, and hypertension.

Table 4 shows the results of the study variables' independency. A model of multiple regression was built with BBS values as the dependent variable and entered all variables that are associated with BBS with p<0.10. Only antimalarial use, handgrip, and tobacco exposure remained as independent variables.

Table 4 presents the results of the multivariable regression analysis. A multiple linear regression model was constructed using the BBS score as the dependent variable, with all variables showing an association with the BBS score at p<0.10 entered into the model. After adjustment, only antimalarial use, handgrip strength, and tobacco exposure remained independently associated with BBS scores.

## DISCUSSION

This study found that Brazilian individuals with RA had poorer postural balance and experienced more falls than healthy controls. In this cohort, postural imbalance was independently associated with reduced handgrip strength—a surrogate marker of overall muscle function—antimalarial use, and tobacco exposure.

**Table 1.** Sample description and controls data pairing.

| | Rheumatoid arthritis | Controls | p-value |
|---|---|---|---|
| Number (n) | 101 | 94 | |
| Female sex—n (%) | 90 (89.1) | 76 (80.8) | 0.11 |
| Age-years—median (IQR) | 58 (52–65) | 55 (47–5) | 0.10 |
| Tobacco exposure (*) | 41 (40.5) | 15 (15.9) | 0.0001 |
| BBS values—median (IQR) | 55 (53–56) | 56 (56–56) | 0.0001 |
| BBS<45 (n) | 6 (5.9%) | 0 | 0.02 |
| At least one fall last year (n) | 12 (11.8%) | 2 (2.1%) | 0.01 |
| Hand grip strength values—median (IQR) | 18.0 (12.5–21.0) | 22.0 (18.0–30.0) | <0.0001 |
| Positive rheumatoid factor—n (%) | 70/93 (75.2) | | |
| Disease duration—years—median (IQR) | 12.0 (8.0–17.0) | | |
| Treatment—n (%) | | | |
| Antimalarials | 7 (6.9) | | |
| Methotrexate | 35 (34.6) | | |
| Leflunomide | 46 (45.5) | | |
| Anti-TNF | 19 (18.8) | | |
| Anti-IL-6 | 4 (3.9) | | |
| Rituximab | 4 (3.9) | | |
| iJAK | 22 (21.7) | | |
| Prednisone | 30 (29.7) | | |
| Inflammatory indexes—median (IQR) | | | |
| DAS 28 ESR | 2.8 (2.22–3.58) | | |
| CDAI | 4.0 (0–10.0) | | |
| SDAI | 4.1 (0.74–10.7) | | |
| ESR—mm | 32.0 (21.0–45.0) | | |
| CRP—mg/dL | 12.0 (8.0–18.5) | | |
| Comorbidities—n (%) | | | |
| Arterial hypertension | 49 (48.5) | | |
| Diabetes mellitus | 10 (9.9) | | |

n: number; IQR: interquartile range; BBS: Berg Balance Scale; TNF: tumor necrosis factor; iJAK: Janus kinase inhibitors, ESR: erythrocyte sedimentation rate; CRP: C-reactive protein; DAS-28: disease activity score using 28 joints; CDAI: clinical disease activity index; SDAI: simplified disease activity index. *Current and former smokers.

**Table 2.** Correlation studies of Berg Balance Scale with numerical variables.

| Variable | R | 95%CI | p-value |
|---|---|---|---|
| Hand grip values | 0.29 | 0.10 to 0.47 | 0.002 |
| Age | -0.41 | -0.56 to -0.22 | <0.0001 |
| DAS 28 ESR | -0.06 | -0.26 to 0.14 | 0.55 |
| CDAI | -0.06 | -0.26 to 0.14 | 0.53 |
| SDAI | -0.09 | -0.30 to 0.12 | 0.40 |
| ESR | -0.12 | -0.31 to 0.08 | 0.22 |
| C-reactive protein | 0.05 | -0.14 to 0.25 | 0.57 |
| Disease duration | -0.17 | -0.36 to 0.02 | 0.07 |

DAS-28 ESR: disease activity score using 28 joints and ESR; ESR: erythrocyte sedimentation rate; CDAI: clinical disease activity index; SDAI: simplified disease activity index.

Handgrip strength is widely recognized as a simple and cost-effective tool for assessing global muscle strength, overall health status, and physical activity levels[11,12]. It has been extensively used as a predictor of sarcopenia[12]. Recent evidence emphasizes that weakness in the lower extremities constitutes a critical risk factor for falls, highlighting the importance of lower-limb muscle strength and quality in maintaining postural control and preventing imbalance. Notably, the primary hip abductors—the gluteus maximus, medius, and minimus—play an essential role in effective recovery of postural stability[13].

According to Torii et al.[14], sarcopenia affects approximately one-third of patients with RA, compared to 10% in the general population. Unlike primary sarcopenia, typically seen in the

**Table 3.** Study of the Berg Balance Scale according to nominal variables.

| Variable | Median BBS values in patients with the variable | Median BBS values in patients without the variable | p-value |
|---|---|---|---|
| Female sex | 55.0 (52.7–56.0) | 56.0 (53.0–56.0) | 0.10 |
| Tobacco exposure | 54.0 (52.0–56.0) | 55.0 (54.0–56.0) | 0.008 |
| Rheumatoid factor | 55.0 (52.7–56.0) | 54.0 (51.5–55.0) | 0.35 |
| Medications used (*) | | | |
| Antimalarials | 53.0 (47.0–54.0) | 55.0 (53.0–56.0) | 0.02 |
| Methotrexate | 56.0 (54.0–56.0) | 54.0 (52.0–56.0) | 0.0007 |
| Leflunomide | 55.0 (52.0–56.0) | 55.0 (53.0–56.0) | 0.90 |
| Anti-TNF | 55.0 (54.0–56.0) | 55.0 (52.0–56.0) | 0.48 |
| Jak inhibitors | 55.0 (52.7–56.0) | 55.0 (53.0–56.0) | 0.61 |
| Prednisone | 55.0 (52.7–56.0) | 55.0 (53.0–56.0) | 0.52 |
| Antidepressants/anxiolytics | 54.0 (52.0–56.0) | 55.0 (53.0–56.0) | 0.06 |
| Comorbidities | | | |
| Arterial hypertension | 54.0 (52.0–56.0) | 56.0 (53.2–56.0) | 0.02 |
| Diabetes | 54.0 (49.5–55.0) | 55.0 (53.0–56.0) | 0.13 |

(*) - anti-IL-6 and rituximab use not studied due to small sample. BBS: Berg Balance Scale; TNF: tumor necrosis factor.

**Table 4.** Univariate and multivariable linear regression analyses of factors associated with Berg Balance Scale scores in patients with rheumatoid arthritis.

| Variable | Univariate analysis β (95%CI) | p | Multivariable analysis β (95%CI) | p-value |
|---|---|---|---|---|
| Age | -0.086 (-0.149 to -0.022) | 0.009 | -0.023 (-0.090 to 0.045) | 0.506 |
| Handgrip strength | 0.124 (0.045 to 0.202) | 0.002 | 0.094 (0.018 to 0.171) | 0.016 |
| Disease duration | -0.075 (-0.140 to -0.011) | 0.023 | -0.046 (-0.110 to 0.018) | 0.154 |
| Antimalarial use | -2.512 (-5.008 to -0.016) | 0.049 | -2.754 (-5.121 to -0.387) | 0.023 |
| Methotrexate use | 2.019 (0.721 to 3.316) | 0.003 | 0.799 (-0.527 to 2.125) | 0.235 |
| Tobacco exposure | -1.748 (-3.018 to -0.478) | 0.008 | -1.916 (-3.208 to -0.624) | 0.004 |

Dependent variable: Berg Balance Scale (BBS) score. **β**: regression coefficient; CI: confidence interval; CNS: central nervous system. Variables entered into the multivariable model were sex, age, disease duration, hypertension, handgrip strength, methotrexate use, antimalarial use, tobacco exposure, and use of antidepressants/anxiolytics. The multivariable model included all variables selected a priori based on clinical relevance and/or univariate analysis.

elderly, RA-related sarcopenia may occur in younger adults and is characterized by loss of skeletal muscle mass with preserved or increased fat mass[15], differing from the generalized weight loss of cachexia or heart failure[16]. Its pathogenesis involves immune dysregulation, hormonal imbalance, and chronic cytokine-mediated inflammation (TNF-α, IL-1, IL-6, IFN-γ), which promote proteolysis and impaired muscle regeneration[17]. Despite therapeutic advances in inflammation control and joint protection, interventions targeting muscle recovery remain limited. No pharmacological therapy has been approved for sarcopenia; however, resistance training shows strong evidence for improving muscle strength and mass[18].

In this study, antimalarial use was also associated with postural imbalance. Although rarely emphasized, ototoxicity is a recognized adverse effect of these drugs. Chloroquine-induced ototoxicity tends to appear acutely, whereas hydroxychloroquine-related effects emerge with prolonged use[19]. According to Muller et al., the inner ear's melanocytes, crucial for cochleovestibular function, may undergo structural and functional damage, and melanin loss can lead to reactive oxygen species accumulation, which is toxic to cochlear hair cells[20]. Clinical consequences include sensorineural hearing loss, tinnitus, and imbalance[21]. The association observed in the present study between postural imbalance and antimalarial use requires confirmation in larger, independent cohorts. Furthermore, investigations that consider cumulative drug exposure may provide additional insights into this relationship. Unlike systemic lupus erythematosus, in which antimalarials are a cornerstone of therapy, RA can usually be managed with several alternative disease-modifying agents. Accordingly, this issue warrants further investigation and clarification.

Tobacco exposure remained independently associated with poorer balance performance in this cohort. In a study of 4,224 Japanese men, smokers performed significantly worse on one-leg balance tests than nonsmokers[22]. Similarly, forest workers who smoked showed greater balance impairment[23]. Nicotine has deleterious effects on motor, sensory, cognitive, and attentional functions, weakens postural muscles, reduces inner ear perfusion, and disrupts vestibular function[24]. Additionally, tobacco use worsens clinical outcomes in RA, potentially impairing lower-limb joint function—an essential component of postural control[25]. These combined effects suggest that smoking may influence both systemic and vestibular pathways implicated in postural control, potentially contributing to balance disturbances and offering a plausible explanation for the observed association.

In agreement with previous studies, RA patients in our study experienced significantly more falls than controls (OR 6.2). This finding raises clinical concern given that RA is frequently associated with reduced bone mass and increased fracture risk due to both disease-related bone loss and prolonged glucocorticoid use. Osteoporotic fractures, particularly hip fractures, are a major cause of morbidity and mortality worldwide[4].

The present study has several limitations, including a relatively small sample size and the predominance of participants with low disease activity, which limited a more comprehensive assessment of the impact of systemic inflammation. In addition, fall history was based on self-reported data from the preceding year and may therefore be subject to recall bias. The absence of a detailed clinical evaluation of lower-limb joint status further restricted a deeper understanding of the mechanisms underlying postural imbalance in RA. Despite these constraints, the findings may offer clinically meaningful insights by indicating that antimalarial use, muscle weakness, and tobacco exposure could be associated with impaired balance in Brazilian patients with RA. Although exploratory, these findings highlight the importance of addressing potentially modifiable risk factors during routine clinical care. A multidisciplinary approach integrating optimal inflammatory control, structured resistance exercise, nutritional optimization, and smoking cessation may help reduce sarcopenia, improve postural balance, and ultimately enhance functional outcomes in patients with RA.

## CONCLUSION

These findings suggest that postural imbalance is a clinically relevant feature of RA that may involve factors beyond inflammatory disease activity alone. The independent associations with reduced muscle strength and tobacco exposure suggest that neuromuscular function and smoking-related factors may contribute substantially to impaired balance, highlighting opportunities for targeted rehabilitation and preventive interventions. The association observed with antimalarial use warrants further investigation to clarify whether it reflects a true treatment effect, disease-related characteristics, or residual confounding. Routine assessment of postural balance may help identify patients at increased risk of falls and support multidisciplinary interventions aimed at preserving mobility, reducing fall risk, and maintaining functional independence.

## AUTHORS' CONTRIBUTIONS

**ALT:** Data curation, Investigation, Methodology, Writing – original draft. **BF:** Data curation, Investigation, Methodology, Writing – original draft. **CMM:** Data curation, Investigation, Methodology, Writing – original draft. **NAH:** Data curation,

Investigation, Methodology, Writing – original draft. **TLS:** Methodology, Supervision, Writing – review & editing. **RN:** Methodology, Supervision, Writing – review & editing. **BSK:** Methodology, Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Wiegmann S, Armbrecht G, Borucki D, Buehring B, Buttgereit F, Detzer C, et al. Balance and prospective falls in patients with rheumatoid arthritis. BMC Musculoskelet Disord. 2022;23(1):549. https://doi.org/10.1186/s12891-022-05489-1
2. Pizzigalli L, Micheletti Cremasco M, Mulasso A, Rainoldi A. The contribution of postural balance analysis in older adult fallers: a narrative review. J Bodyw Mov Ther. 2016;20(2):409-17. https://doi.org/10.1016/j.jbmt.2015.12.008
3. Lyon A, Lester S, Stanhope J, Lynch T, Black R, Barrett C, et al. Dual-energy X-ray absorptiometry and anti-osteoporotic medication use in Australian patients with early rheumatoid arthritis using data from the Australian Rheumatology Association Database. Intern Med J. 2025;55(7):1127-35. https://doi.org/10.1111/imj.70065
4. Furuya T, Inoue E, Yamanaka H, Harigai M, Tanaka E. Increasing fracture incidence over 13 years in patients with rheumatoid arthritis from the IORRA cohort. J Bone Miner Metab. 2025;43(4):411-8. https://doi.org/10.1007/s00774-025-01598-8
5. Guo X, Pei J, Wei Y, Zhang G, Yan F, Han L. Prevalence and risk factors of falls in adults with rheumatoid arthritis: a systematic review and meta-analysis. Semin Arthritis Rheum. 2023;60:152186. https://doi.org/10.1016/j.semarthrit.2023.152186
6. Furuya T, Yamagiwa K, Ikai T, Inoue E, Taniguchi A, Momohara S, et al. Associated factors for falls and fear of falling in Japanese patients with rheumatoid arthritis. Clin Rheumatol. 2009;28(11):1325-30. https://doi.org/10.1007/s10067-009-1229-5
7. Aletaha D, Neogi T, Silman AJ, Funovits J, Felson DT, Bingham CO, et al. 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. Arthritis Rheum. 2010;62(9):2569-81. https://doi.org/10.1002/art.27584
8. Medeiros MM, Oliveira BM, Cerqueira JV, Quixadá RT, Oliveira ÍM. Correlation of rheumatoid arthritis activity indexes (Disease Activity Score 28 measured with ESR and CRP, Simplified Disease Activity Index and Clinical Disease Activity Index) and agreement of disease activity states with various cut-off points in a Northeastern Brazilian population. Rev Bras Reumatol. 2015;55(6):477-84. https://doi.org/10.1016/j.rbr.2014.12.005
9. Berg KO, Wood-Dauphinee SL, Williams JI, Maki B. Measuring balance in the elderly: validation of an instrument. Can J Public Health. 1992;83(Suppl 2):S7-11. PMID: 1468055
10. Shumway-Cook A, Baldwin M, Polissar NL, Gruber W. Predicting the probability for falls in community-dwelling older adults. Phys Ther. 1997;77(8):812-9. https://doi.org/10.1093/ptj/77.8.812
11. Leong DP, Teo KK, Rangarajan S, Lopez-Jaramillo P, Avezum A, Orlandini A, et al. Prognostic value of grip strength: findings from the Prospective Urban Rural Epidemiology (PURE) study. Lancet. 2015;386(9990):266-73. https://doi.org/10.1016/S0140-6736(14)62000-6
12. Zaccagni L, Toselli S, Bramanti B, Gualdi-Russo E, Mongillo J, Rinaldo N. Handgrip strength in young adults: association with anthropometric variables and laterality. Int J Environ Res Public Health. 2020;17(12):4273. https://doi.org/10.3390/ijerph17124273
13. Friedman B, Beamer BA, Beans J, Gray V, Alon G, Ryan A, et al. Neuromuscular electrical stimulation to maximize hip abductor strength and reduce fall risk in older veterans: protocol for a randomized controlled trial. JMIR Res Protoc. 2025;14:e68082. https://doi.org/10.2196/68082
14. Torii M, Itaya T, Minamino H, Katsushima M, Fujita Y, Tanaka H, et al. Management of sarcopenia in patients with rheumatoid arthritis. Mod Rheumatol. 2023;33(3):435-40. https://doi.org/10.1093/mr/roac095
15. Bennett JL, Pratt AG, Dodds R, Sayer AA, Isaacs JD. Rheumatoid sarcopenia: loss of skeletal muscle strength and mass in rheumatoid arthritis. Nat Rev Rheumatol. 2023;19(4):239-51. https://doi.org/10.1038/s41584-023-00921-9
16. Letarouilly JG, Flipo RM, Cortet B, Tournadre A, Paccou J. Body composition in patients with rheumatoid arthritis: a narrative literature review. Ther Adv Musculoskelet Dis. 2021;13:1759720X211015006. https://doi.org/10.1177/1759720X211015006
17. Papamatthaiou S, Menelaou P, El Achab Oussallam B, Moschou D. Recent advances in bio-microsystem integration and Lab-on-PCB technology. Microsyst Nanoeng. 2025;11(1):78. https://doi.org/10.1038/s41378-025-00940-4
18. Jia PW, Lin JZ, Zou YW, Ouyang ZM, Yang Y, Yang KM, et al. The characteristic of muscle function for sarcopenia in patients with rheumatoid arthritis: a large-scale real-world cross-sectional study. Medicina (Kaunas). 2025;61(4):551. https://doi.org/10.3390/medicina61040551
19. Reynard P, Thai-Van H. Drug-induced hearing loss: listening to the latest advances. Therapie. 2024;79(2):283-95. https://doi.org/10.1016/j.therap.2023.10.011
20. Muller R. Systemic toxicity of chloroquine and hydroxychloroquine: prevalence, mechanisms, risk factors, prognostic and screening possibilities. Rheumatol Int. 2021;41(7):1189-202. https://doi.org/10.1007/s00296-021-04868-6
21. Bortoli R, Santiago M. Chloroquine ototoxicity. Clin Rheumatol. 2007;26(11):1809-10. https://doi.org/10.1007/s10067-007-0662-6
22. Saito T, Miyatake N, Nishii K. Relationship between cigarette smoking and one leg with eyes closed balance in Japanese men. Environ Health Prev Med. 2015;20(5):388-91. https://doi.org/10.1007/s12199-015-0457-3
23. Iki M, Ishizaki H, Aalto H, Starck J, Pyykkö I. Smoking habits and postural stability. Am J Otolaryngol. 1994;15(2):124-8. https://doi.org/10.1016/0196-0709(94)90061-2
24. Rana Y, Vaish H. Evaluation of the static and dynamic balance in single and dual tasks among active smokers and non-smokers. Healthcare Low-Resource Settings. 2023;11(1). https://doi.10.4081/hls.2023.11159
25. Toprak CŞ, Duruöz MT, Gündüz OH. Static and dynamic balance disorders in patients with rheumatoid arthritis and relationships with lower extremity function and deformities: a prospective controlled study. Arch Rheumatol. 2018;33(3):328-34. https://doi.org/10.5606/ArchRheumatol.2018.6720

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260435

# Investigation of the relationship between expression patterns of plasma microRNAs and pregnancy complications

Nevin Karakus[1]*, Elif Donmezoglu Kutukcu[1], Sumeyye Yildirim[1]

## SUMMARY

**OBJECTIVE:** The aim of the study was to determine the association between plasma expression levels of microRNAs (miRNAs; MIR126-3p; and MIR375-3p) and pregnancy complications.

**METHODS:** This prospective study included 39 pregnant women who admitted to Obstetrics and Gynecology Polyclinic of Tokat Gaziosmanpasa University, Faculty of Medicine. 28 of these pregnant women were healthy, while 11 had pregnancy complications (five hypertensive disorders of pregnancy and six gestational diabetes). Hypertensive disorders include preeclampsia, gestational hypertension, and chronic hypertension. The study included pregnant women older than 18 years of age with singleton pregnancies and no history of cancer. Quantitative real time-polymerase chain reaction was used to determine the expression patterns of MIR126-3p and MIR375-3p in maternal plasma. Fold change ($2^{-\Delta\Delta Ct}$) analysis was used to examine the microRNA expression patterns.

**RESULTS:** MIR375-3p expression levels demonstrated a statistically significant difference between two groups. It was significantly lower in women with pregnancy complications than healthy pregnant women.

**CONCLUSION:** We found that MIR375-3p might be a useful indicator for non-invasive diagnosis and play a crucial part in predicting pregnancy complications. Given the limited patient cohort, more research using bigger samples is required to validate our findings.

**KEYWORDS:** Pregnancy complications. MicroRNAs. Real-time polymerase chain reaction.

## INTRODUCTION

The most important problems facing contemporary obstetrics are pregnancy complications, which impact at least 15% of pregnancies globally and include preeclampsia, gestational diabetes, and premature delivery[1]. They are linked to long-term negative outcomes for both the mother and the child, and they have a significant influence on maternal and fetal death and morbidity. Nonetheless, it is generally acknowledged that early detection of women at risk may enhance successful therapies, therefore lowering the prevalence of these potentially fatal disorders and the expense of newborn care.

MicroRNAs (miRNAs) are small non-coding RNAs that attach to mRNA, prevent translation, or encourage degradation in order to control gene expression. Since their discovery in 1993, miRNAs have been linked to essential cellular processes, including tissue development, proliferation, differentiation, apoptosis, and maintenance of homeostasis[2]. To prevent destruction, tissue miRNAs are released into the bloodstream, coupled to certain proteins or high-density lipoproteins, or encased in exosomes, microvesicles, or apoptotic bodies[2]. We refer to them as circulating miRNAs. These miRNAs play roles in vascular health, coagulation control, and immune tolerance during pregnancy.

The pregnancy complications have been linked to altered expression levels of circulating miRNAs that could be helpful in diagnosis and prognosis of these diseases[3,4]. Certain miRNAs carried by microvesicles have the ability to alter endothelial cell activity, which can affect vascular health and placental function[5-10]. Also, when compared to the normal placenta, the preeclampsia placenta exhibits abnormal expression of many miRNAs: some upregulated (MIR210 and MIR141) and some downregulated (MIR223 and MIR376c)[11-14].

MIR126, a highly conserved miRNA, seems to be one of the most crucial miRNAs for healthy development. During pregnancy, MIR126, a well-known regulator of angiogenesis and endothelial function, is essential for preserving vascular integrity by limiting leukocyte adherence to the endothelium[15]. There have been reports of both increased and decreased MIR126 expression in placentas and serums of preeclampsia patients[15-18]. There are also studies that found

[1]Tokat Gaziosmanpasa University, Faculty of Medicine, Department of Medical Biology – Tokat, Turkey.
*Corresponding author: nevinbalci@hotmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 25, 2026. Accepted on July 07, 2026.
Scientific Editor: José Maria Soares Júnior

increased expression of MIR375-3p in the plasma of preeclampsia patients[19,20].

Because miRNAs in maternal blood are stable, they are perfect for non-invasive diagnostic testing in cases of pregnancy complications. Clinicians can learn more about the placental health of pregnant women and perhaps enable early intervention by examining their levels and activity[21]. The aim of the current study was to examine the association between the plasma expression levels of miRNAs (MIR126-3p and MIR375-3p) and pregnancy complications.

## METHODS

### Study samples

The ethical suitability of this study was confirmed by the Non-Intervention Scientific Research Ethics Committee of Tokat Gaziosmanpasa University Faculty of Medicine in April 8, 2025 (Registration No: 25-MOBAEK-116) and all the participants signed informed consent form for this study. This study included 39 pregnant women who visited the Obstetrics and Gynecology Polyclinic of Tokat Gaziosmanpasa University Faculty of Medicine. 28 of these pregnant women were healthy, while 11 had pregnancy complications (five hypertensive disorders of pregnancy and six gestational diabetes). Hypertensive disorders include preeclampsia, gestational hypertension, and chronic hypertension. After 20 weeks of pregnancy, new-onset hypertension (diastolic blood pressure [DBP] ≥90 mm Hg/systolic blood pressure [SBP] ≥140 mm Hg) is considered gestational hypertension. Women were presumed to have chronic hypertension if they had hypertension prior to 20 weeks of pregnancy. New-onset hypertension (DBP ≥90 mm Hg/SBP ≥140 mm Hg) with proteinuria or signs of end-organ damage after 20 weeks of pregnancy was considered preeclampsia. Gestational diabetes was diagnosed using a single-step approach. For this, a 75-g oral glucose tolerance test was performed in the morning after at least 8 h of fasting. The presence of at least one of the following: plasma glucose (fasting) values: ≥92 mg/dL, 1-h ≥180 mg/dL, 2-h ≥153 mg/dL was considered sufficient for a gestational diabetes diagnosis. The inclusion criteria for all the pregnant women was being older than 18 years of age with singleton pregnancies and having no history of cancer. Other inclusion criteria included having a hypertensive disorder for the patient group and not having any pregnancy complications for the control group. Exclusion criteria for all pregnant women include having intrauterine infection or any other surgical, medical or other obstetric complaint, a history of cancer, multiple pregnancy, and being under 18 years of age. For this study, a total of 38 samples were detected by using the G*Power 3.1.9.7 program with 78% power, 5% margin of error, and an effect size of 0.8.

### Sample preparation and the use of real-time polymerase chain reaction analysis

Five milliliters of maternal blood samples were placed in ethylenediaminetetraacetic acid tubes, which were centrifuged for 10 min at 3,000 rpm. Following centrifugation, the plasma was transposed to microcentrifuge tubes and kept at -80°C until it was required. The miRNAs in the plasma samples were extracted by following the manufacturer's instructions for the SanPrep Column microRNA Mini-Preps Kit (Code: SK8811, Biobasic, Canada). The complementary DNA (cDNA) conversion for both miRNAs was carried out using the miR-cDNA Synthesis kit found in the single-miRNA quantitative polymerase chain reaction (qPCR) test kit (Cat No: Q07-01-10, Atlas Biotechnologies, Türkiye), which was present independently for both miRNAs. The expression patterns of MIR375-3p and MIR126-3p in plasma were next ascertained by a RT-PCR using the obtained cDNAs and the miR-qPCR MasterMix that came with the single-miRNA qPCR test kit (Cat. No.: Q07-01-10, Atlas Biotechnologies, Türkiye). For reverse transcription and RT-PCR, respectively, the miR-cDNA Synthesis kit and miR-qPCR MasterMix include the following essential ingredients: RNase inhibitor, reverse transcriptase, antibody-mediated hot-start Taq DNA polymerase, miRNA-specific stem-loop primer, forward and reverse primers, $MgCl_2$, nucleoside triphosphates (dNTPs), ROX reference dye, SBYR green dye, stabilizer, and enhancer. An Applied Biosystems 7500 Fast RT-PCR System (Applied Biosystems, Foster City, CA, USA) was the device that was used to apply RT-PCR. The comparative threshold cycle (Ct) method was used to evaluate each miRNA's relative expression. The calculations directly used the Ct values of the target genes and the reference gene (RNU6-1), which was used for normalization. Fold change ($2^{-\Delta\Delta Ct}$) analysis was applied to analyze the miRNA expression patterns.

### Statistical analysis

The statistical analysis was carried out using IBM Statistical Package for the Social Sciences Statistical Package version 27.0. Characteristics of the samples were stated as either the mean ± standard deviation (SD) or the number, depending on availability. To assess the statistical value of circulating miRNAs, a one-way analysis of variance (ANOVA) test was employed. Different pregnancy complications and healthy pregnancies were compared pairwise using Tukey's test and one-way ANOVA. Using the boxplot method, the graphs were created. Statistical significance was defined as a two-tailed p-value of less than 0.05.

## RESULTS

This study was conducted with a total of 39 pregnant women, including 28 healthy pregnant women and 11 pregnant women with pregnancy complications. Of the 11 pregnant women with pregnancy complications, 6 had gestational diabetes, and 5 had hypertensive disorders (Table 1). The age and trimester distributions of pregnant women with pregnancy complications and healthy pregnant women were consistent (Table 1). Furthermore, no difference was seen between the two groups for parity, gravity, and abortion numbers. However, the SBP and DBP values differed between the two groups. SBP and DBP values were significantly elevated in pregnant women with pregnancy complications compared to healthy pregnant women as expected (Table 1).

According to statistical analyses, while the expression levels of MIR126-3p did not differ between healthy pregnant women and pregnant women with pregnancy complications, the expression levels of MIR375-3p showed a statistically significant variation between these two groups (Table 2). The expression level of MIR375-3p was lower in pregnant women with pregnancy complications compared to healthy pregnant women. When pregnant women with pregnancy complications were separated into categories based on disease type, such as gestational diabetes and hypertensive disorders of pregnancy, and compared among themselves and with healthy pregnant women, no statistically significant association was observed between expression levels of both miRNAs and these categories (Table 3).

## DISCUSSION

In this study examining the relationship between pregnancy complications and miRNA expression, a statistically significant

**Table 1.** The demographical and clinical characteristics of pregnant women with pregnancy complications and healthy pregnant women.

| Characteristics | Pregnant women with pregnancy complications (n=11) | Healthy pregnant women (n=28) | p-value |
|---|---|---|---|
| Age, years | 30.96±1.347 | 32.70±4.165 | 0.056 |
| Gestational age, weeks | 29.18±5.250 | 31.64±7.041 | 0.302 |
| Trimester, n (%) | | | |
| 2. Trimester (15–27 weeks) | 4 (36.4) | 8 (28.6) | 0.411 |
| 3. Trimester (28–33 weeks) | 4 (36.4) | 6 (21.4) | |
| 3. Trimester (34–40 weeks) | 3 (27.3) | 14 (50.0) | |
| Complication, n (%) | | | |
| Hypertensive disorders of pregnancy | 5 (45.5) | - | - |
| Gestational diabetes | 6 (54.5) | - | |
| SBP, mmHg | 129.43±18.329 | 116.44±7.649 | **0.008** |
| DBP, mmHg | 80.429±10.014 | 72.72±7.525 | **0.033** |
| Gravida, n | 2.89±0.601 | 3.12±1.705 | 0.701 |
| Parity, n | 1.44±0.527 | 1.58±1.391 | 0.784 |
| Abortion, n | 0.44±0.726 | 0.54±0.905 | 0.780 |

Mean plus standard deviation values are presented for all the characteristics. BMI: body mass index; DBP: diastolic blood pressure; SBP: systolic blood pressure. The statistically significant values are represented in bold.

**Table 2.** Expression levels and p-values of MIR375-3p and MIR126-3p in pregnant women with pregnancy complications and healthy pregnant women.

| miRNA | Group | n | Expression level ($2^{-\Delta\Delta Ct}$) | p-value |
|---|---|---|---|---|
| MIR375-3p | Pregnant women with pregnancy complications | 11 | 0.52±0.210 | **0.015** |
| | Healthy pregnant women | 28 | 1.56±1.321 | |
| MIR126-3p | Pregnant women with pregnancy complications | 11 | 2.89±1.250 | 0.403 |
| | Healthy pregnant women | 28 | 2.14±2.829 | |

miRNA: microRNA. The statistically significant values are represented in bold.

**Table 3.** Expression levels and p-values of MIR375-3p and MIR126-3p in pregnancies with different complications and healthy pregnancy.

| Group | n | Expression level ($2^{-\Delta\Delta Ct}$) | p-value | | |
|---|---|---|---|---|---|
| | | MIR375-3p | Hypertensive disorders of pregnancy versus Pregnancy diabetes | Hypertensive disorders of pregnancy versus Healthy pregnancy | Pregnancy diabetes versus Healthy pregnancy |
| Hypertensive disorders of pregnancy | 5 | 0.46±0.205 | 0.986 | 0.136 | 0.153 |
| Gestational diabetes | 6 | 0.58±0.218 | | | |
| Healthy pregnancy | 28 | 1.56±1.321 | | | |
| | | MIR126-3p | Hypertensive disorders of pregnancy versus Pregnancy diabetes | Hypertensive disorders of pregnancy versus Healthy pregnancy | Pregnancy diabetes versus Healthy pregnancy |
| Hypertensive disorders of pregnancy | 5 | 3.45±1.525 | 0.781 | 0.537 | 0.965 |
| Gestational diabetes | 6 | 2.43±0.830 | | | |
| Healthy pregnancy | 28 | 2.14±2.829 | | | |

association was found between expression levels of MIR375-3p and pregnancy complications. However, when we divided our patient group into two categories based on disease type, such as gestational diabetes and hypertensive disorders of pregnancy, this association disappeared. The possible explanation for this is that the number of patients in both categories decreased as a result of this division.

Widespread vasculogenesis, pseudovasculogenesis, and angiogenesis are all carried out by the growing placenta to aid in the development of the fetus. These mechanisms are mediated by angiogenic factors and trophoblast invasion. Preeclampsia, one of the hypertensive disorders of pregnancy, is thought to be caused by early failure of these mechanisms during pregnancy[22]. Angiogenesis signaling and endothelial cell vascular integrity are primarily promoted by MIR126. There have been reports of both increased and downregulated MIR126 expression in placentas and serums of preeclampsia patients[15-18]. However, in our study, we found no association, neither downregulation nor upregulation, between pregnancy complications, including preeclampsia, and MIR126 expression.

In recent years, MIR375 has been widely reported in a range of immunological activities, indicating that it may be an important regulator of cell differentiation and an integral part of immune system regulation[23]. There were two reports that examined the association between preeclampsia and MIR375 expression[19,20]. Both studies found increased expression of MIR375-3p in the plasma of preeclampsia patients. These findings differ from ours. In our findings, MIR375 expression was decreased in pregnant women with pregnancy complications, but this association disappeared when we grouped the patients as having gestational diabetes and hypertensive disorder of pregnancy. This may be explained by variances in the populations under study and sample size as well as changes in pre-analytical and analytical procedures. Furthermore, because miRNAs are involved in epigenetic processes, additional variables including maternal age, lifestyle, illnesses, or drug misuse may have an impact on their expression levels. Thus, it is important to look into the effects of patient-to-patient variation in plasma miRNA levels.

The primary constraint of the current study was the limited sample size. In all, 39 samples were examined. Larger sample numbers are essential for identifying subtle but physiologically meaningful changes in miRNA research. Therefore, it's possible that this study's limited sample size prevented it from finding correlations found in earlier research. Furthermore, the restricted statistical power resulting from the small sample size may be the reason why there was no difference in the expression patterns of these miRNAs between the hypertensive diseases of pregnancy and gestational diabetes. Even evaluating miRNA expression levels by itself is insufficient to comprehend the intricacy of pregnancy complications.

In conclusion, we found a statistically significant association between expression levels of circulating MIR375-3p and pregnancy complications. These findings should be validated in larger, multi-center cohorts. Most significantly, validation is necessary in connection to clinical symptoms in order to evaluate the clinical importance of circulating miRNAs in pregnancy complications and to ascertain sensitivity, specificity, and predictive value. To enhance health care for all pregnant women in the future, it is likely that a panel of several biomarkers, including genomic, transcriptomic, proteomic, and biophysical ones, paired with maternal features, will be required due to the multifactorial nature of these illnesses.

## ETHICS APPROVAL

The ethical suitability of this study was confirmed by the Non-Intervention Scientific Research Ethics Committee of Tokat Gaziosmanpasa University Faculty of Medicine in April 8, 2025 (Registration No: 25-MOBAEK-116).

## AUTHORS' CONTRIBUTIONS

**NK:** Conceptualization, Formal Analysis, Investigation, Resources, Validation, Writing – original draft. **EDK:** Data curation, Investigation, Methodology, Resources, Writing – review & editing. **SY:** Data curation, Investigation, Methodology, Resources, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Wallis AB, Saftlas AF, Hsia J, Atrash HK. Secular trends in the rates of preeclampsia, eclampsia, and gestational hypertension, United States, 1987-2004. Am J Hypertens. 2008;21(5):521-6. https://doi.org/10.1038/ajh.2008.20
2. Lycoudi A, Mavreli D, Mavrou A, Papantoniou N, Kolialexi A. miRNAs in pregnancy-related complications. Expert Rev Mol Diagn. 2015;15(8):999-1010. https://doi.org/10.1586/14737159.2015.1053468
3. Konečná B, Tóthová Ľ, Repiská G. Exosomes-associated DNA-new marker in pregnancy complications? Int J Mol Sci. 2019;20(12):2890. https://doi.org/10.3390/ijms20122890
4. Bardan CR, Ioniță I, Iordache M, Călămar-Popovici D, Todorescu V, Popescu R, et al. Epigenetic biomarkers in thrombophilia-related pregnancy complications: mechanisms, diagnostic potential, and therapeutic implications: a narrative review. Int J Mol Sci. 2024;25(24):13634. https://doi.org/10.3390/ijms252413634
5. Morelli AE, Sadovsky Y. Extracellular vesicles and immune response during pregnancy: a balancing act. Immunol Rev. 2022;308(1):105-22. https://doi.org/10.1111/imr.13074
6. Martin C, Bergamelli M, Malnou CE, D'Angelo G. Placental extracellular vesicles in maternal-fetal communication during pregnancy. Biochem Soc Trans. 2022;50(6):1785-95. https://doi.org/10.1042/BST20220734
7. Maligianni I, Yapijakis C, Nousia K, Bacopoulou F, Chrousos GP. Exosomes and exosomal non-coding RNAs throughout human gestation (Review). Exp Ther Med. 2022;24(3):582. https://doi.org/10.3892/etm.2022.11518
8. Aharon A, Rebibo-Sabbah A, Ahmad RS, Dangot A, Bar-Lev TH, Brenner B, et al. Associations of maternal and placental extracellular vesicle miRNA with preeclampsia. Front Cell Dev Biol. 2023;11:1080419. https://doi.org/10.3389/fcell.2023.1080419
9. Kumar MA, Baba SK, Sadida HQ, Marzooqi SA, Jerobin J, Altemani FH, et al. Extracellular vesicles as tools and targets in therapy for diseases. Signal Transduct Target Ther. 2024;9(1):27. https://doi.org/10.1038/s41392-024-01735-1
10. Coenen CS, Hidalgo TN, Lynn T, Jones DM, Salmon JE, Chamley LW, et al. Antiphospholipid-exposed trophoblast-derived extracellular vesicles express elevated levels of TLR7/8-activating microRNAs and induce endometrial endothelial activation, in part, through TLR7. J Reprod Immunol. 2024;164:104255. https://doi.org/10.1016/j.jri.2024.104255
11. Fu G, Ye G, Nadeem L, Ji L, Manchanda T, Wang Y, et al. MicroRNA-376c impairs transforming growth factor-β and nodal signaling to promote trophoblast cell proliferation and invasion. Hypertension. 2013;61(4):864-72. https://doi.org/10.1161/HYPERTENSIONAHA.111.203489
12. Ospina-Prieto S, Chaiwangyen W, Herrmann J, Groten T, Schleussner E, Markert UR, et al. MicroRNA-141 is upregulated in preeclamptic placentae and regulates trophoblast invasion and intercellular communication. Transl Res. 2016;172:61-72. https://doi.org/10.1016/j.trsl.2016.02.012
13. Jairajpuri DS, Malalla ZH, Sarray S, Mahmood N. Analysis of differential expression of hypoxia-inducible microRNA-210 gene targets in mild and severe preeclamptic patients. Noncoding RNA Res. 2021;6(1):51-7. https://doi.org/10.1016/j.ncrna.2021.03.001
14. Liu X, Li Z, Lu D. MicroRNA-223-3p downregulates the inflammatory response in preeclampsia placenta via targeting NLRP3. BMC Pregnancy Childbirth. 2024;24(1):175. https://doi.org/10.1186/s12884-024-06371-9
15. Chu X, Gu Y, Sheng W, Sun J, Morgan JA, Lewis DF, et al. Downregulation of miR-126-3p expression contributes to increased inflammatory response in placental trophoblasts in preeclampsia. J Reprod Immunol. 2021;144:103281. https://doi.org/10.1016/j.jri.2021.103281
16. Jaszczuk I, Koczkodaj D, Kondracka A, Kwaśniewska A, Winkler I, Filip A. The role of miRNA-210 in pre-eclampsia development. Ann Med. 2022;54(1):1350-6. https://doi.org/10.1080/07853890.2022.2071459
17. Soobryan N, Kumar A, Moodley J, Mackraj I. The role and expression of pro/antiangiogenic factors and microRNAs in gestational hypertension and pre-eclampsia. Eur J Obstet Gynecol Reprod Biol. 2023;290:38-42. https://doi.org/10.1016/j.ejogrb.2023.09.012

18. Ozler S, Kebapcilar A, Ozdemir EM, Mert M, Arıkan MN, Celik C. Are vascular endothelium and angiogenesis effective MicroRNA biomarkers associated with the prediction of early-onset preeclampsia (EOPE) and adverse perinatal outcomes? Reprod Sci. 2024;31(3):803-10. https://doi.org/10.1007/s43032-023-01367-w

19. Sandrim VC, Eleuterio N, Pilan E, Tanus-Santos JE, Fernandes K, Cavalli R. Plasma levels of increased miR-195-5p correlates with the sFLT-1 levels in preeclampsia. Hypertens Pregnancy. 2016;35(2):150-8. https://doi.org/10.3109/10641955.2015.1122034

20. Akgör U, Ayaz L, Çayan F. Expression levels of maternal plasma microRNAs in preeclamptic pregnancies. J Obstet Gynaecol. 2021;41(6):910-4. https://doi.org/10.1080/01443615.2020.1820465

21. Ghosh S, Thamotharan S, Fong J, Lei MYY, Janzen C, Devaskar SU. Circulating extracellular vesicular microRNA signatures in early gestation show an association with subsequent clinical features of pre-eclampsia. Sci Rep. 2024;14(1):16770. https://doi.org/10.1038/s41598-024-64057-w

22. Wang S, Olson EN. AngiomiRs--key regulators of angiogenesis. Curr Opin Genet Dev. 2009;19(3):205-11. https://doi.org/10.1016/j.gde.2009.04.002

23. Liu Y, Wang Q, Wen J, Wu Y, Man C. MiR-375: a novel multifunctional regulator. Life Sci. 2021;275:119323. https://doi.org/10.1016/j.lfs.2021.119323

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260456

# Functional shoulder performance and fatigue in young adults with forward head posture with or without scapular dyskinesis

Seda Bicici Ulusahin[1]*, Husna Balci[2], Nur Barlas Yalcin[1], Melike Nur Ozcan[1], Zeynep Berfu Ecemiș[3], Savas Kudas[4]

## SUMMARY

**OBJECTIVE:** The aim of this study was to compare functional shoulder performance and fatigue among young adults with forward head posture, forward head posture with scapular dyskinesis (forward head posture+scapular dyskinesia), and healthy controls.

**METHODS:** This cross-sectional study included 85 young adults aged 18–25 years, allocated to three groups: forward head posture+scapular dyskinesia (n=25), forward head posture (n=30), and control (n=30). Forward head posture was assessed using the craniovertebral angle, scapular dyskinesia with the Lateral Scapular Slide Test, functional shoulder performance with Timed Functional Arm and Shoulder Test, and fatigue with the Chalder Fatigue Scale.

**RESULTS:** Significant between-group differences were found for dominant-side, non-dominant-side, and total Timed Functional Arm and Shoulder Test scores ($p<0.001$). The forward head posture+scapular dyskinesia group had lower dominant-side and total Timed Functional Arm and Shoulder Test scores than the both forward head posture and control groups ($p=0.040$, $p<0.001$, and $p=0.031$, $p<0.001$, respectively), while both postural disorder groups had lower non-dominant-side Timed Functional Arm and Shoulder Test scores than controls (forward head posture+scapular dyskinesia: $p<0.001$; forward head posture: $p=0.003$). Total and physical fatigue scores were higher in both the forward head posture+scapular dyskinesia and forward head posture groups than in controls (total fatigue: $p=0.010$ and $p=0.011$; physical fatigue: both $p=0.019$), whereas mental fatigue did not differ significantly among the groups ($p=0.163$).

**CONCLUSION:** Functional shoulder performance was lower in the forward head posture and forward head posture+scapular dyskinesia groups than in the controls, with the lowest scores in the forward head posture+scapular dyskinesia group. Total and physical fatigue were higher in both postural disorder groups, whereas mental fatigue did not differ between groups.

**KEYWORDS:** Posture. Scapula. Shoulder joint. Fatigue. Young adult.

## INTRODUCTION

Forward head posture (FHP) is a common postural deviation characterized by the head being positioned forward relative to the neck in the sagittal plane[1]. FHP is characterized by excessive activity and shortening of the levator scapulae, sternocleidomastoid, upper trapezius, and suboccipital muscles. Decreased activity and muscle weakness also occur in stabilizing muscle groups such as the deep neck flexors, lower trapezius, rhomboids, and serratus anterior[2]. An imbalance between the superficial and deep muscles can lead to increased compensatory muscle activity. This can result in early fatigue and proprioceptive deficiencies[3]. FHP and scapular dyskinesia (SD) are frequently seen together, but the cause-and-effect relationship between the two situations remains unclear[4,5]. The scapula acts as a functional bridge between the neck and shoulder regions. It regulates the force transfer during glenohumeral and scapulothoracic movements, maintaining stability and mobility of the upper extremity[6-8].

It has been reported that the coexistence of FHP and SD exacerbates the fatigue and neck and shoulder discomfort[9]. Previous studies have mainly examined FHP and SD in relation to pain, muscle performance, range of motion, and shoulder function, both individually and in combination[10,11]. However, evidence on performance-based shoulder function and perceived fatigue remains limited. Therefore, the aim of this study is to examine the effects of FHP alone and in combination with SD on functional shoulder performance and perception of general fatigue, compared with healthy individuals.

[1]University of Health Sciences, Gulhane Faculty of Physiotherapy and Rehabilitation, Department of Physiotherapy and Rehabilitation – Ankara, Turkey.
[2]Cappadocia University, School of Health Sciences, Department of Physiotherapy and Rehabilitation – Nevșehir, Turkey.
[3]Cukurova University, Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation – Adana, Turkey.
[4]Acıbadem Hospital, Department of Sports Medicine – Ankara, Turkey.
*Corresponding author: seda.ulusahin@sbu.edu.tr
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 20, 2026. Accepted on July 04, 2026.
Scientific Editor: José Maria Soares Júnior

We hypothesized that individuals with both FHP and SD would have lower functional shoulder performance and higher perceived fatigue compared with individuals with isolated FHP and with a healthy control group.

# METHODS

## Participants

All participants received information about the study and provided written informed consent. Inclusion criteria were age between 18 and 25 years, no history of neck, shoulder or back pain/injury, and full active shoulder range of motion. Exclusion criteria were previous orthopedic surgery involving the back, shoulder, or neck, signs of cervical radiculopathy, rheumatic disease, or pain during testing. Participants were divided into three groups: FHP, FHP+SD, and control. The study was conducted at the Gulhane Faculty of Physiotherapy and Rehabilitation Laboratory between October 2025 and January 2026, with ethical approval from the Scientific Research Ethics Committee of the University of Health Sciences (decision no: 2024-493), in accordance with the Declaration of Helsinki.

## Assessments

### *Forward head posture*

FHP was assessed using the craniovertebral angle (CVA). Standardized lateral photographs were taken with participants standing relaxed and looking forward. Adhesive markers were placed on the tragus and C7 spinous process. A smartphone was mounted on a tripod at shoulder height, 1 m from the participant. Images were analyzed in Kinovea software (version 2025.1.1) by a different blinded clinician. CVA was defined as the angle between the C7–tragus line and the horizontal reference line; CVA<48° indicated FHP[12].

### *Scapular dyskinesia*

The Lateral Scapular Slide Test was performed by a blinded clinician with 18 years of experience and no role in other assessments. In three positions, the distance between the inferior angle of the scapula and the corresponding thoracic spinous process parallel to it was measured bilaterally with a tape measure. The positions were arms relaxed at the sides, hands on hips with 45° shoulder abduction, and 90° shoulder abduction with maximum internal rotation. One measurement was recorded per position, and a side-to-side difference >1.5 cm in any position was considered positive[13].

### *Functional shoulder performance test*

It was assessed using the Timed Functional Arm and Shoulder Test (TFAST), as described by Shah et al[14]. TFAST was administered by two assessors blinded to group allocation and uninvolved in clinical classification; one counted repetitions and the other timed the test. The test includes three tasks: hand-to-head and back touch, wall wash, and gallon jug lift. After a standardized demonstration, participants performed each task as quickly as possible for 30 s. In the hand-to-head and back touch task, participants stood upright with the arm at the side of the body and were asked to touch the back of the head. In the wall wash task, participants moved the hand in circular motions over a 30-cm pattern marked on the wall, in both inward and outward directions. In the gallon jug lift task, a 3.78-kg plastic jug was lifted from a counter to a shelf and returned. All tasks were performed bilaterally, with a 2-min rest between tasks, and the total score=[Hand-to-head and back+(Wall wash inward+outward)+Gallon jug lift][14,15].

### *Fatigue*

The Chalder Fatigue Scale was used to assess fatigue. Data were collected via Google Forms in the laboratory setting on a tablet before clinical assessments. The scale consists of 11 items. The first seven items assess physical fatigue, and the remaining four items assess mental fatigue. Each item is rated on a four-point Likert scale. The total score ranges from 0 to 33. The higher scores indicate greater levels of fatigue[16].

## Statistical analysis

Data were analyzed using International Business Machines Statistical Package for the Social Sciences Statistics 29.0. Descriptive statistics were presented as mean ± standard deviation (SD). Normality was assessed using Shapiro-Wilk, histograms, and Levene's test. Group comparisons were performed using one-way analysis of variance (ANOVA). Because six separate one-way ANOVA models were conducted for the TFAST and Chalder Fatigue Scale outcomes, Holm-Bonferroni correction was applied to the omnibus ANOVA p-values to control the family-wise type I error rate across outcome measures. For outcomes that remained significant after this correction, Tukey honestly significant difference (HSD) post-hoc tests were used. Categorical variables were compared using chi-square test.

Sample size was determined by an a priori power analysis using G*Power software (version 3.1.9.7). Based on the study by Shah et al., which established the sensitivity of the TFAST in detecting differences in functional performance, a large effect size (f=0.40) was anticipated[14]. With an alpha level of 0.05 and a desired power of 0.80, the minimum required sample size for

three groups was calculated as 66. Allowing for a 10% dropout rate, a minimum of 25 participants per group was planned. The final sample consisted of 85 participants distributed across three groups (Figure 1. Study Flow). A post-power analysis was also performed with G*power. The analysis showed that the power of the study was 99% [Power (1 - β err prob)], Critical F=3.11.

## RESULTS

A total of 85 participants were included in the study: 25 in the FHP+SD group, 30 in the FHP group, and 30 in the control group. Age, sex distribution, and body mass index were similar across the groups (p=0.322, p=0.494, and p=0.220, respectively) (Table 1). After Holm-Bonferroni correction across the six omnibus ANOVA tests, significant group effects remained for dominant-side TFAST (p<0.001), non-dominant-side TFAST (p<0.001), and total TFAST scores (p<0.001), total fatigue (p=0.011), and physical fatigue (p=0.015). The group effect for mental fatigue was not significant (p=0.163).

For dominant-side TFAST, the FHP+SD group had lower scores than both the FHP groups (p=0.040) and the control group (p<0.001); the FHP group had lower scores than the control group (p=0.040). For TFAST on the non-dominant side, both the FHP+SD groups (p<0.001) and the FHP group (p=0.003) had lower scores than the control group, while the FHP+SD versus FHP comparison did not reach statistical significance (p=0.056). For total TFAST, the FHP+SD group had lower scores than the both FHP groups (p=0.031) and the control group (p<0.001); the FHP group also had lower scores than the control group (p=0.007). Both the FHP+SD and FHP groups had higher total fatigue scores than the control group (p=0.010 and p=0.011, respectively), with no significant difference between the two postural disorder groups (p=0.982). Similarly, both the FHP+SD and FHP groups had higher physical fatigue scores than the control group (both p=0.019), whereas the difference between the FHP+SD and FHP groups was not significant (p=0.990) (Table 2).

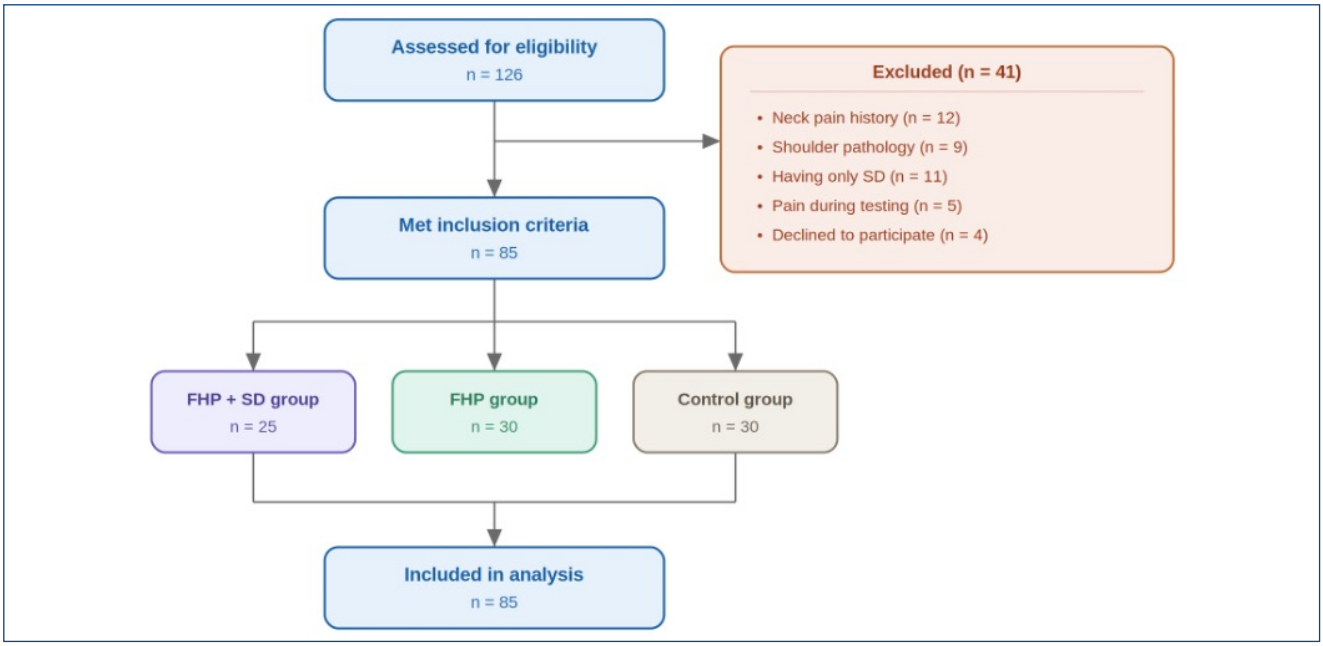



**Figure 1.** Flow diagram of participant recruitment and group classification.

**Table 1.** Demographic characteristics of the study groups.

| Variable | FHP+SD (n=25) | FHP (n=30) | Control (n=30) | p-value |
|---|---|---|---|---|
| Age (years), mean±SD | 21.20±1.94 | 21.17±1.97 | 20.53±1.76 | 0.322 |
| Sex (female/male), n (%) | 20/5 (80/20) | 27/3 (90/10) | 24/6 (80/20) | 0.494 |
| BMI (kg/m$^2$), mean±SD | 23.98±3.89 | 24.45±4.99 | 22.65±3.17 | 0.220 |

SD: standard deviation; BMI: body mass index; FHP: forward head posture; SD: scapular dyskinesia. Between-group differences assessed by one-way ANOVA (continuous variables) and chi-square test (sex).

**Table 2.** Between-group comparisons of functional shoulder performance and fatigue.

| Assessment parameters, mean ± SD | FHP+SD n=25 | FHP n=30 | Control n=30 | p | Holm-adjusted p | Post-hoc Tukey HSD |
|---|---|---|---|---|---|---|
| TFAST score, dominant shoulder | 100.40±10.84 | 108.83±13.92 | 116.87±12.46 | <0.001 | <0.001 | 1<2 p=0.040<br>1<3 p<0.001<br>2<3 p=0.040 |
| TFAST score, non-dominant shoulder | 93.56±10.63 | 100.67±11.58 | 110.60±11.29 | <0.001 | <0.001 | 1<3 p<0.001<br>2<3, p=0.003<br>1 vs. 2, p=0.056 |
| TFAST score, total score | 193.96±19.83 | 209.50±24.02 | 227.47±22.33 | <0.001 | <0.001 | 1<2 p=0.031<br>1<3 p<0.001<br>2<3 p=0.007 |
| Chalder Fatigue Scale, total score | 14.32±4.96 | 14.10±4.53 | 10.63±4.14 | 0.004 | 0.011 | 1>3 p=0.010<br>2>3 p=0.010<br>1 vs. 2 p=0.982 |
| Chalder fatigue, Physical Fatigue Subscale | 9.72±3.26 | 9.60±3.55 | 7.23±3.14 | 0.008 | 0.015 | 1>3 p=0.019<br>2>3 p=0.019<br>1 vs. 2, p=0.990 |
| Chalder fatigue, Mental Fatigue Subscale | 4.60±2.24 | 4.50±2.10 | 3.63±1.97 | 0.163 | 0.163 | Not interpreted because omnibus ANOVA was not significant |

SD: standard deviation; TFAST: Timed Functional Arm and Shoulder Test; FHP: forward head posture; SD: scapular dyskinesia; ANOVA: analysis of variance; Tukey HSD: Tukey's honestly significant difference test. Post-hoc: Tukey HSD. Values are presented as mean±standard deviation. Group codes: 1=FHP+SD, 2=FHP, 3=Control. Six separate one-way ANOVA models were performed; therefore, Holm-Bonferroni correction was applied to the omnibus ANOVA p-values to control the family-wise type I error rate across outcomes. Tukey HSD post-hoc tests were used for pairwise comparisons within each ANOVA and were interpreted only for outcomes with significant Holm-adjusted omnibus ANOVA results.

## DISCUSSION

In this study, individuals with forward head posture, those with both forward head posture and scapular dyskinesis, and a control group without postural deviations were compared with respect to functional shoulder performance and fatigue. According to the findings, TFAST scores were significantly lower in the FHP and FHP+SD groups than in the control group. This decrease was particularly pronounced in individuals with both FHP and SD. Furthermore, total fatigue and physical fatigue scores were significantly higher in these groups than in the control group.

To our knowledge, no study has directly compared individuals with only FHP, individuals with both FHP and SD, and healthy controls in terms of functional shoulder performance, nor has any study quantitatively demonstrated a decrease in performance. However, the literature supports that the combination of FHP and SD can negatively affect shoulder biomechanics and function by increasing muscle imbalances and postural impairment[9,11,17]. Because TFAST evaluates range of motion, endurance, and strength together, it is a suitable tool for identifying this combined effect[14]. Furthermore, it has been reported that a difference of more than 21 repetitions in this test is considered clinically significant[15]. The total TFAST difference between the FHP+SD and control groups (193.96 vs. 227.47; difference=33.51 repetitions) exceeded this threshold; however, as this value was derived from patients with shoulder problems, its relevance to asymptomatic young adults is uncertain and should be interpreted cautiously. When SD accompanies FHP, scapulothoracic mechanics may be further impaired: the scapula may not provide stable support for glenohumeral movement, the acromiohumeral distance may decrease, and trapezius-serratus anterior force balance may change[18-20]. Taken together, the lower dominant-side and total TFAST scores in the FHP+SD group suggest poorer shoulder function, although the non-dominant-side difference between FHP+SD and FHP did not reach statistical significance.

The significantly higher total and physical fatigue scores observed in both the FHP and FHP+SD groups, compared with the control group, suggest that these postural deviations may be particularly related to perceived physical fatigue. This finding is consistent with the literature reporting that forward head displacement increases the mechanical load on the cervical extensor muscles, leading to compensatory activation of the upper trapezius and levator scapulae muscles and increasing physical fatigue[2,21]. However, the lack of a significant difference between the FHP and FHP+SD

groups for total and physical fatigue should be interpreted cautiously. Fatigue perception is subjective, and the absence of objective measures such as electromyography may have limited our ability to detect subtle differences between these groups. Future studies should use different fatigue scales together with objective methods to better clarify fatigue-related changes. The lack of a significant difference between the groups in mental fatigue may be due to our sample consisting of young adults aged 18–25 years. Relatively high cognitive capacity in this age group and the absence of a significant mental burden from postural compensation mechanisms may explain this situation.

This study has several limitations. First, physical activity level and screen time were not recorded; therefore, their potential confounding effects on shoulder performance and fatigue cannot be excluded. Future studies should measure and control for these variables. Second, the inclusion of only young adults may limit the generalizability of the findings to other age groups, where the cumulative effects of postural disorders on shoulder performance and fatigue may differ. Future studies should therefore include a broader age range. Although the lateral scapular slide test was applied by a single experienced clinician using a standardized clinical protocol, intra- or inter-rater reliability was not assessed in this study. Therefore, possible measurement error or misclassification of scapular dyskinesis cannot be fully excluded.

## CONCLUSION

In summary, individuals with FHP and FHP+SD exhibited significantly lower functional shoulder performance than the control group, and this impairment was more pronounced in the FHP+SD group. Total and physical fatigue scores were significantly higher in both the FHP and FHP+SD groups than in the control group, with no significant difference between the two groups. Mental fatigue scores did not differ significantly among the groups.

## AUTHORS' CONTRIBUTIONS

**SBU:** Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **HB:** Data curation, Formal Analysis, Investigation, Writing – original draft. **NBY:** Conceptualization, Data curation, Investigation, Methodology, Writing – original draft. **MNO:** Data curation, Formal Analysis, Investigation, Writing – review & editing. **ZBE:** Data curation, Formal Analysis, Investigation, Writing – original draft. **SK:** Data curation, Methodology, Supervision, Writing – original draft, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Sheikhhoseini R, Shahrbanian S, Sayyadi P, O'Sullivan K. Effectiveness of therapeutic exercise on forward head posture: a systematic review and meta-analysis. J Manipulative Physiol Ther. 2018;41(6):530-9. https://doi.org/10.1016/j.jmpt.2018.02.002
2. Lee H, Lee Y. Effects of postural changes using a standing desk on the craniovertebral angle, muscle fatigue, work performance, and discomfort in individuals with a forward head posture. Healthcare (Basel). 2024;12(23):2436. https://doi.org/10.3390/healthcare12232436
3. Demirtaş G, Güzel NA. Forward head posture: examination from biomechanical, postural, and therapeutic perspectives-current literature and rehabilitation methods. Laodicea J. Rehab Sci. 2025;3(1):10-7. https://doi.org/10.5281/zenodo.15277916
4. Vongsirinavarat M, Wangbunkhong S, Sakulsriprasert P, Petviset H. Prevalence of scapular dyskinesis in office workers with neck and scapular pain. Int J Occup Saf Ergon. 2023;29(1):50-5. https://doi.org/10.1080/10803548.2021.2018855
5. Singla D, Veqar Z. Association between forward head, rounded shoulders, and increased thoracic kyphosis: a review of the literature. J Chiropr Med. 2017;16(3):220-9. https://doi.org/10.1016/j.jcm.2017.03.004
6. Karaağaç A, Arslan SA, Keskin ED. Assessment of pain, scapulothoracic muscle strength, endurance and scapular dyskinesis in individuals with and without nonspecific chronic neck pain: a cross-sectional study. J Bodyw Mov Ther. 2023;35:261-7. https://doi.org/10.1016/j.jbmt.2023.04.008
7. Longo UG, Risi Ambrogioni L, Berton A, Candela V, Massaroni C, Carnevale A, et al. Scapular dyskinesis: from basic science to ultimate treatment. Int J Environ Res Public Health. 2020;17(8):2974. https://doi.org/10.3390/ijerph17082974
8. Lefèvre-Colau MM, Nguyen C, Palazzo C, Srour F, Paris G, Vuillemin V, et al. Recent advances in kinematics of the shoulder complex in healthy people. Ann Phys Rehabil Med. 2018;61(1):56-9. https://doi.org/10.1016/j.rehab.2017.09.001
9. Javdaneh N, Barati A, Shojaedin S, Mirzaei RM. Is chronic neck pain related to scapular dyskinesia? A systematic review. BMC Musculoskelet Disord. 2025;26(1):585. https://doi.org/10.1186/s12891-025-08916-1
10. Mahmoud NF, Hassan KA, Abdelmajeed SF, Moustafa IM, Silva AG. The relationship between forward head posture and neck pain: a systematic review and meta-analysis. Curr Rev Musculoskelet Med. 2019;12(4):562-77. https://doi.org/10.1007/s12178-019-09594-y
11. Panagiotopoulos AC, Crowther IM. Scapular Dyskinesia, the forgotten culprit of shoulder pain and how to rehabilitate. SICOT J. 2019;5:29. https://doi.org/10.1051/sicotj/2019029
12. Zárate-Tejero CA, Rodríguez-Rubio PR, Brandt L, Krauss J, Hernández-Secorún M, Hidalgo-García C, et al. Measuring craniovertebral angle reference values in adults using Kinovea software. Appl Sci (Basel). 2024;14(19):8639. https://doi.org/10.3390/app14198639

13. Odom CJ, Taylor AB, Hurd CE, Denegar CR. Measurement of scapular asymetry and assessment of shoulder dysfunction using the lateral scapular slide test: a reliability and validity study. Phys Ther. 2001;81(2):799-809. https://doi.org/10.1093/ptj/81.2.799

14. Shah KM, Baker T, Dingle A, Hansmeier T, Jimenez M, Lopez S, et al. Early development and reliability of the timed functional arm and shoulder test. J Orthop Sports Phys Ther. 2017;47(6):420-31. https://doi.org/10.2519/jospt.2017.7136

15. Safford DW, Shah KM, Turner JA, McClure PW. Measurement properties of the timed functional arm and shoulder test in patients with shoulder problems. Phys Ther. 2023;103(2):pzac144. https://doi.org/10.1093/ptj/pzac144

16. Adın RM, Ceren AN, Salcı Y, Fil Balkan A, Armutlu K, Ayhan Kuru Ç. Dimensionality, psychometric properties, and population-based norms of the Turkish version of the Chalder Fatigue Scale among adults. Health Qual Life Outcomes. 2022;20(1):161. https://doi.org/10.1186/s12955-022-02074-x

17. Castelein B, Cools A, Parlevliet T, Cagnie B. Are chronic neck pain, scapular dyskinesis and altered scapulothoracic muscle activity interrelated?: a case-control study with surface and fine-wire EMG. J Electromyogr Kinesiol. 2016;31:136-43. https://doi.org/10.1016/j.jelekin.2016.10.008

18. Dehqan B, Delkhoush CT, Mirmohammadkhani M, Ehsani F. Does forward head posture change subacromial space in active or passive arm elevation? J Man Manip Ther. 2021;29(4):227-34. https://doi.org/10.1080/10669817.2020.1854010

19. Lewis JS, Green A, Wright C. Subacromial impingement syndrome: the role of posture and muscle imbalance. J Shoulder Elbow Surg. 2005;14(4):385-92. https://doi.org/10.1016/j.jse.2004.08.007

20. Thigpen CA, Padua DA, Michener LA, Guskiewicz K, Giuliani C, Keener JD, et al. Head and shoulder posture affect scapular mechanics and muscle activity in overhead tasks. J Electromyogr Kinesiol. 2010;20(4):701-9. https://doi.org/10.1016/j.jelekin.2009.12.003

21. Lotfian S, Fesharaki MJ, Shahabbaspour Z, Akbarzadeh H, Moezy A. The impact of forward head posture on neck muscle endurance and thickness in women with chronic neck pain: a cross-sectional study. BMC Musculoskelet Disord. 2025;26(1):468. https://doi.org/10.1186/s12891-025-08705-w

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260490

# Thyroid function and anemia: results from the Brazilian Longitudinal Study of Adult Health

Alinne de Santo Nascimento[1], Marcos Rafael Nogueira Cavalcante[1], Angelita Gomes de Souza[1], Isabela Judith Martins Benseñor[1]*

## SUMMARY

**OBJECTIVE:** The aim of this study was to analyze the association of thyroid stimulating hormone and free-thyroxin with anemia as well as the role of several sociodemographic variables in these associations.

**METHODS:** This is a cross-sectional analysis of the Brazilian Longitudinal Study of Adult Health study (2008–2010), a prospective cohort study of civil servants in six locations in Brazil. Inclusion criteria: to be an active or retired civil servant of one of the institutions and 35–74 years at enrollment. Exclusion criteria: pregnancy less than 4 months before entering the study for women, intention to quit the institution soon, live outside the metropolitan region where the investigation center is located or communication problems. Anemia was defined as hemoglobin < 12 g% in women and < 13 g% in men. After exclusions, 13,247 participants were included in the analysis. Models were presented without adjustment and adjusted for sociodemographic and clinical risk factors. The study was approved by the Ethic Committee.

**RESULTS:** Median age was 51.9 (51–53 years), 54% of women There was an association of less presence of anemia with higher free-thyroxin levels considering the entire sample: OR 0.80 (95%CI 0.65–0.98) for the second quintile; OR 0.55 (95%CI 0.43–0.70) for third quintile; OR 0.49 (95%CI 0.38–0.63) for fourth quintile; and OR 0.52 (95%CI 0.40–0.66) for fifth quintile for free-thyroxin levels. This association was also present when analyzing the sample by sex.

**CONCLUSION:** Higher levels of free-thyroxin are associated with lower odds of anemia. No associations were found between thyroid stimulating hormone levels and anemia. In analysis, according to sex, No differences were found between men and women.

**KEYWORDS:** Thyroid hormones. Thyroid stimulating hormone (TSH). Anemia.

## INTRODUCTION

Thyroid diseases are very prevalent worldwide with an age-standardized prevalence around 5%[1]. In Brazil, prevalence of subclinical and overt hypothyroidism were respectively 5.4 and 7.4%[2]. Anemia is also a prevalent disease, impacting 27% of the world's population[3]. Some studies reported an association of clinical and subclinical hypothyroidism with anemia, as patients with hypothyroidism may present low antioxidant potential, leading to an increase in inflammatory cytokines that induce alterations, potentially interfering with different erythropoiesis pathways, which could lead to anemia[3-5].

Some studies evaluated the association of thyroid diseases and anemia. In 2018, a meta-analysis conducted as part of the Thyroid Studies Collaboration investigated the associations between thyroid dysfunction and anemia. The findings showed that levels of hemoglobin were lower across all thyroid dysfunction subgroups compared to euthyroid participants[6]. In the cross-sectional study of Vliet et al[7], including 445,482 participants, the authors did not report any association between thyroid stimulating hormone (TSH) or free-thyroxin (FT4) levels with anemia. In another cross-sectional analysis (n=12,337), the adjusted Hb was 0.22 g/dL lower in overt hypothyroidism compared to euthyroid participants[4]. In a prospective analysis, patients with subclinical and overt hypothyroidism at baseline showed a trend of increased risk of developing anemia during follow-up[8]. Puy et al conducted a study to assess whether levothyroxine treatment leads to a rise in hemoglobin levels in patients with subclinical hypothyroidism. The change in hemoglobin levels was not significantly different between the levothyroxine and the placebo groups[9].

In this study, we aim to analyze the association of TSH and FT4 with anemia using data from the Brazilian Longitudinal

[1]Universidade de São Paulo, Hospital Universitário, Centro de Pesquisa Clínica e Epidemiológica – São Paulo (SP), Brazil.
*Corresponding author: isabensenor@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on April 22, 2026. Accepted on July 23, 2026.
Scientific Editor: Roseli Nomura

Study of Adult Health (ELSA-Brasil), as well as the role of several sociodemographic variables in this association.

## METHODS

This is a cross-sectional analysis using data about thyroid function and hemoglobin levels from ELSA-Brasil baseline (2008–2010), a cohort study following 15,105 civil servants aged 35–74 years from six institutions (Federal Universities of Minas Gerais, Espírito Santo, Rio Grande do Sul and Bahia, Oswaldo Cruz Foundation and University of São Paulo), with the objective of investigating the incidence of cardiovascular disease, diabetes and other chronic diseases. In this analysis we used data from the baseline examination (2008–2010)[10]. Inclusion criteria of the study were to be an active or retired civil servant of one of the institutions and 35–74 years at enrollment. Exclusion criteria were pregnancy less than 4 months before entering the study for women, intention to quit the institution soon, live outside the metropolitan region where the investigation center is located or communication problems. More information about the study is elsewhere[10-12].

### Study population

Of 15,105 participants included in the baseline of ELSA-Brasil, 53 presented missing data on anemia, 1,604 presented missing data on thyroid function and 254 for using medication that interfered with the thyroid function, totaling after exclusions a sample of 13,247.

### Thyroid function

TSH and FT4 were determined by a third-generation immunoenzymatic assay (Roche Diagnostic, Manheim, Germany) at baseline. The analysis included euthyroid participants (TSH 0.4–4.0 mIU/L) with no use of levothyroxine or antithyroid medication and participants with subclinical hypothyroidism (TSH>4.0 mIU/L, FT4 within reference levels) and overt hypothyroidism (TSH>4.0 mIU/L, low FT4). TSH reference range was 0.4–4.0 mIU/L, FT4 reference range was 0.93–1.7 ng/dL.

### Anemia

Blood cell counts were performed with automated equipment and anemia was defined in women if Hb<12 g/dL and in men if Hb<13 g/dL.

### Other variables

Sociodemographic variables were categorized as: age, included as a continuous variable; race, classified according to the Brazilian National Census as White, Mixed (pardo), Black, Asian, and indigenous; Education was categorized as less than high-school, high school and incomplete college or complete college or more. Average familiar monthly net income was classified as: <1,245, 1,245–3,319 and ≥3,320 (1US$=2R$ 2.00 at baseline). Comorbidities are defined as the presence of at least one hypertension, diabetes, dyslipidemia, depression, anxiety, obesity, coronary heart diseases, and stroke. It was categorized as no comorbidity, 1 comorbidity and two or more comorbidities.

### Statistical analysis

Categorical variables were expressed as counts and percentages and compared using the chi-squared test. The Shapiro-Wilk test was applied to the age, TSH, and FT4 values to assess the assumption of normality. As they presented a non-normal distribution, we used the Kruskal-Wallis test to compare continuous variables.

TSH and FT4 were divided into quintiles. We used logistic regression models expressed as odds ratio (OR) and their respective 95%CIs presented in models without adjustment and multivariable adjusted for both sociodemographic and chronic conditions (Model 1): for age, sex, race, education, hypertension, and diabetes. These analyses were performed with "R" software version 4.4.2 and R Studio 2024.04.2 Build 764.

## RESULTS

Table 1 shows general characteristics of the sample according to FT4 quintiles. Mean hemoglobin levels were 13.8 (1.4) for the first quintile; 14.0 (1.3) for the second quintile; 14.2 (1.3) for the third quintile; 14.3 (1.3) for the fourth quintile; 14.4 (1.3) for the fifth quintile (p<0.0001). Table 2 shows the general characteristic of the sample according to TSH quintiles.

Frequency of anemia decreased according to FT4 quintiles. The lower FT4 quintile presented the higher frequency of women (p<0.0001). Frequency of White, Asian and Indigenous participants increased according to FT4 quintiles, while frequency of Mixed and Black participants decreased (p<0.0001). Frequency of College education (p<0.0001) was higher in participants of the first quintile. Frequency of hypertension (p<0.0001) was higher in participants of the fifth quintile. Although differences in body mass index (BMI) were statistically significant according to FT4 levels, those differences were clinically irrelevant.

Table 3 shows the association of anemia with TSH and FT4 in the sample according to sex and the entire sample. Using the first quintile as reference, there was an association of less presence of anemia with higher FT4 levels present in males, and also in females. Such association was also present when considering the entire sample.

## DISCUSSION

A systematic review conducted by Floriani et al[8] following 12,337 participants with available baseline TSH, FT4, and hemoglobin levels used multiple linear regression analysis to compare hemoglobin levels across TSH categories. They found that mean hemoglobin levels in patients with subclinical hypothyroidism were lower than in euthyroid individuals at baseline, but without clinical relevance because of the small difference. In our sample the number of participants with anemia was too small to permit analysis according to hypothyroidism or euthyroid participants, which restricted our analysis to the association of anemia with TSH and FT4 levels. We reported an association of increasing FT4 levels with lower odds of anemia (p<0.0001).

In the analysis of Wopereis et al[6], TSH, FT4, and hemoglobin data of 42,162 participants, with 51.5% women, from 16 different cohorts were analyzed, of whom 459 (1.1%) had overt hypothyroidism, 2,930 (6.9%) had subclinical hypothyroidism, and 36,081 (85.6%) were euthyroid. Of those, 4,274 (10.1%) had anemia at baseline, 15.9% in the overt hypothyroid group and 11.6% in the subclinical hypothyroid group. They found no statistically significant differences between men and women when asserting the risk of developing anemia in any of the groups. For females: OR 1.79 (95%CI 1.30–2.47) in the overt hypothyroid group; OR 1.23 (95%CI 0.99–1.52) in the subclinical hypothyroid group. For males: OR 2.45 (95%CI 1.45–4.12) in the overt hypothyroid group; OR 1.27

**Table 1.** General characteristics of the sample according to free-thyroxin quintiles.

| General characteristics (ng/dL) | 0.02–1.04 (2,662) | 1.05–1.14 (2,910) | 1.15–1.22 (2,606) | 1.23–1.32 (2,448) | 1.33–5.80 (2,621) | p-value |
|---|---|---|---|---|---|---|
| Anemia (%) | 217 (8.1) | 187 (6.4) | 111 (4.2) | 90 (3.6) | 98 (3.7) | <0.0001 |
| Hemoglobin levels (g/dL) | 13.7 (12.9–14.7) | 13.9 (13.1–15.0) | 14.1 (13.3–15.1) | 14.3 (13.4–15.2) | 14.4 (13.5–15.3) | <0.0001 |
| Woman (%) | 1,700 (63.8) | 1,673 (57.4) | 1,368 (52.4) | 1,216 (49.6) | 1,258 (48.0) | <0.0001 |
| Age median (1Q–3Q) | 51 (45.0–58.0) | 51 (45.0–58.0) | 51 (45.0–58.0) | 51 (45.0–58.0) | 52 (45.0–59.0) | 0.0025 |
| Race (%) | | | | | | |
| White | 1,351 (51.2) | 1,443 (50.0) | 1,373 (53.4) | 1,305 (54.0) | 1,433 (55.4) | <0.0001 |
| Mixed | 772 (29.2) | 848 (29.4) | 682 (26.5) | 655 (27.1) | 659 (25.5) | |
| Black | 442 (16.7) | 502 (17.4) | 423 (16.4) | 362 (15.0) | 353 (13.6) | |
| Asian | 42 (1.5) | 63 (2.1) | 65 (2.5) | 68 (2.8) | 103 (3.9) | |
| Indigenous | 28 (1.0) | 28 (0.9) | 26 (1.0) | 23 (0.9) | 36 (1.3) | |
| Education (%) | | | | | | |
| <High-school | 342 (12.8) | 363 (12.4) | 306 (11.7) | 317 (12.9) | 359 (13.6) | <0.0001 |
| High school/some college | 1,004 (37.7) | 1,106 (38.0) | 895 (32.8) | 805 (32.8) | 829 (31.6) | |
| College or more | 1,316 (49.4) | 1,441 (49.5) | 1,405 (53.9) | 1,326 (54.1) | 1,433 (54.6) | |
| Income (US$) | | | | | | |
| Up to 1,245 | 764 (28.8) | 785 (27.0) | 643 (24.8) | 638 (26.1) | 632 (24.1) | 0.0003 |
| 1,245–3,319 | 1,174 (44.3) | 1,299 (44.8) | 1,153 (44.5) | 1,063 (43.6) | 1,152 (44.0) | |
| ≥3,320 | 710 (26.8) | 814 (28.0) | 795 (30.6) | 737 (30.2) | 829 (31.7) | |
| Hypertension (%) | 875 (32.9) | 978 (33.6) | 914 (35.0) | 845 (34.5) | 1,043 (39.8) | <0.0001 |
| Diabetes (%) | 387 (14.5) | 424 (14.5) | 398 (15.2) | 399 (16.3) | 464 (17.7) | 0.0059 |
| Dyslipidemia (%) | 1,263 (47.6) | 1,308 (45.0) | 1,174 (45.2) | 1,085 (44.3) | 1,240 (47.4) | 0.0552 |
| (BMI kg/m$^2$) median (1Q–3Q) | 27.2 (24.3–30.5) | 26.5 (23.9–29.8) | 26.2 (23.4–29.4) | 26.0 (23.5–28.8) | 25.8 (23.3–28.9) | <0.0001 |
| Comorbidities | | | | | | |
| 0 | 626 (23.5) | 735 (25.2) | 682 (26.1) | 722 (29.4) | 632 (24.1) | <0.0001 |
| 1 | 890 (33.4) | 1,004 (34.5) | 874 (33.5) | 768 (31.3) | 900 (34.3) | |
| 2 or more | 1,146 (43.0) | 1,171 (40.2) | 1,050 (40.2) | 958 (39.1) | 1,089 (41.5) | |

Anemia was defined as <13 g/dL for men and <12 g/dL for women. Comorbidities have been included as the presence of: HAS, diabetes, dyslipidemia, depression, anxiety, obesity, coronary heart disease (HMPA10 and HMPA45), and stroke.

(95%CI 1.03–1.57) in the subclinical hypothyroid group. These findings are similar to ours since we did not detect differences according to sex.

The analysis by Vliet et al[7] using self-reported data on hypothyroidism of the UK Biobank observed higher odds of anemia for subclinical hypothyroidism, but no evidence

**Table 2.** Comparison between euthyroid and subclinical hypothyroidism samples.

| General characteristics | Euthyroid | Subclinical hypothyroidism | p-value |
|---|---|---|---|
| Anemia (%) | 572 (5.2) | 43 (3.8) | 0.0538 |
| Hemoglobin levels (g/dL) | 14.1 (13.2–15.1) | 14.2 (13.3–15.2) | 0.1011 |
| Women (%) | 5,637 (51.8) | 589 (53.3) | 0.4722 |
| Age, years median (1Q–3Q) | 51 (45–58) | 53 (47–60) | <0.0001 |
| Self-reported race (%) | | | |
| White | 5,475 (50.9) | 667 (60.6) | <0.0001 |
| Mixed | 3,034 (28.2) | 281 (25.5) | |
| Black | 1,842 (17.1) | 107 (9.7) | |
| Asian | 280 (2.6) | 28 (2.5) | |
| Indigenous | 114 (1.0) | 16 (1.4) | |
| Education attainment (%) | | | |
| Less than high-school | 1,399 (12.8) | 156 (14.0) | 0.3287 |
| High school and some college | 3,879 (35.6) | 375 (33.7) | |
| Complete college or more | 5,597 (51.4) | 580 (52.2) | |
| Income (US$) | | | |
| Up to 1,245 | 4,756 (43.9) | 530 (47.8) | 0.0171 |
| 1,245–3,319 | 3,140 (28.9) | 315 (28.4) | |
| ≥3,320 | 2,937 (27.1) | 262 (23.6) | |
| Smoking (%) | | | |
| Never | 6,138 (56.4) | 656 (59.0) | <0.0001 |
| Past | 3,206 (29.4) | 358 (32.2) | |
| Current | 1,530 (14.0) | 97 (8.7) | |
| Alcohol intake (%) | | | |
| Never | 1,124 (10.3) | 122 (11.0) | 0.7825 |
| Past | 2,124 (19.5) | 213 (19.2) | |
| Current | 7,617 (70.1) | 774 (69.7) | |
| Single (%) | 3,572 (32.8) | 395 (35.5) | 0.0729 |
| Hypertension (%) | 3,792 (34.8) | 410 (36.9) | 0.188 |
| Diabetes (%) | 1,692 (15.5) | 167 (15.0) | 0.6824 |
| Dyslipidemia (%) | 4,912 (45.1) | 505 (45.4) | 0.8881 |
| Body-mass index (BMI kg/m$^2$) median (1Q–3Q) | 26.2 (23.6–29.4) | 26.6 (23.9–29.6) | 0.0675 |
| Comorbidities | | | |
| 0 | 2,872 (26.4) | 285 (25.6) | 0.6467 |
| 1 | 3,601 (33.1) | 383 (34.4) | |
| 2 or more | 4,402 (40.4) | 443 (39.8) | |

Anemia was defined as <13 g/dL for men and <12 g/dL for women; Model 1 includes age, race, education attainment, and income; Model 2 includes age, race, education attainment, diabetes mellitus, and hypertension. BMI: body mass index.

of a causal relationship between subclinical hypothyroidism and anemia using euthyroid participants as the reference. It is important to note that the studies of Wopereis et al[6] and Vliet et al[7] analyzed the association of hypothyroidism and anemia while in our analysis we evaluated the association of anemia with TSH and FT4.

In the study of Puy et al[9], 669 participants were divided in two groups for testing whether levothyroxine treatment leads to a rise in hemoglobin levels in people with subclinical hypothyroidism (placebo N=337 and levothyroxine N=332. Median age was 75 years for both groups). At baseline, 13.1% of the participants had anemia in the placebo group compared to 11.4% in the levothyroxine group (mean hemoglobin level was 13.8 g/dL for both groups). TSH levels had no impact on Hb levels after the observation period when patients underwent levothyroxine treatment in any of the subgroups, whether they were related to clinical risk factors such as smoking or alcohol status, or related to comorbidities, such as hypertension or diabetes mellitus.

**Table 3.** Association of thyroid stimulating hormone levels and free-thyroxin with presence of anemia, presented according to men, women, and the entire sample.

| | | TSH | | FT4 |
|---|---|---|---|---|
| Male | Crude model | Q1 | 1 (Reference) | 1 (Reference) |
| | | Q2 | 0.66 (0.39–1.09) | 0.51 (0.33–0.78) |
| | | Q3 | 0.76 (0.47–1.24) | 0.41 (0.26–0.64) |
| | | Q4 | 0.96 (0.60–1.53) | 0.34 (0.21–0.55) |
| | | Q5 | 1.07 (0.68–1.70) | 0.31 (0.19–0.49) |
| | Model 1 | Q1 | 1 (Reference) | 1 (Reference) |
| | | Q2 | 0.66 (0.39–1.11) | 0.52 (0.33–0.79) |
| | | Q3 | 0.79 (0.48–1.30) | 0.43 (0.27–0.68) |
| | | Q4 | 1.04 (0.65–1.67) | 0.38 (0.23–0.62) |
| | | Q5 | 1.14 (0.71–1.83) | 0.31 (0.19–0.51) |
| Female | Crude model | Q1 | 1 (Reference) | 1 (Reference) |
| | | Q2 | 0.80 (0.61–1.05) | 0.94 (0.74–1.18) |
| | | Q3 | 0.88 (0.67–1.14) | 0.60 (0.45–0.79) |
| | | Q4 | 0.80 (0.61–1.05) | 0.54 (0.40–0.72) |
| | | Q5 | 0.76 (0.58–1.00) | 0.58 (0.44–0.78) |
| | Model 1 | Q1 | 1 | 1 |
| | | Q2 | 0.85 (0.64–1.13) | 0.92 (0.72–1.17) |
| | | Q3 | 1.01 (0.76–1.32) | 0.61 (0.46–0.80) |
| | | Q4 | 0.99 (0.74–1.30) | 0.56 (0.41–0.76) |
| | | Q5 | 1.04 (0.78–1.37) | 0.67 (0.50–0.90) |
| The entire sample | Crude model | Q1 | 1 (Reference) | 1 (Reference) |
| | | Q2 | 0.73 (0.57–0.92) | 0.77 (0.63–0.95) |
| | | Q3 | 0.81 (0.64–1.02) | 0.50 (0.39–0.63) |
| | | Q4 | 0.81 (0.64–1.02) | 0.43 (0.33–0.55) |
| | | Q5 | 0.82 (0.65–1.03) | 0.44 (0.34–0.56) |
| | Model 1 | Q1 | 1 (Reference) | 1 (Reference) |
| | | Q2 | 0.81 (0.64–1.04) | 0.80 (0.65–0.98) |
| | | Q3 | 0.95 (0.75–1.21) | 0.55 (0.43–0.70) |
| | | Q4 | 1.01 (0.79–1.28) | 0.49 (0.38–0.63) |
| | | Q5 | 1.06 (0.84–1.35) | 0.52 (0.40–0.66) |

Anemia was defined as <13 g/dL for men and <12 g/dL for women; Model 1 includes age, race, education attainment, diabetes mellitus, and hypertension. TSH: thyroid stimulating hormone; FT4: free-thyroxin.

A previous study in Korea[13] showed an association similar to what was observed in our study. In a sample of 5,352 participants, there was an association of low levels of FT4 with higher frequency of anemia. This association was higher in women than men. In our sample, high levels of FT4 were associated with low odds of anemia both in men and women. Compared to the Korean analysis, in our study there were no differences in the frequency of anemia according to sex. Thyroid hormones have a very important role in hematopoiesis, especially for red blood cell production[14-17], however, the mechanism to explain this association is not clearly understood[16-19]. The results of the study of Kim et al[13] and the findings of the present analysis suggest that FT4 would be more important than TSH in the control of erythropoiesis.

The strengths of this study consists in the large number of participants in a highly mixed sample, setting it apart from previous studies. The multifactorial nature of the cohort allowed us to access a wide variety of variables for cross-sectional analysis. The study has several weaknesses. This is a cross-sectional analysis that can study association but no causality. The percentage of anemia diagnoses according to thyroid function in the total sample, only accounting for a total of 5.3% (N=703) of participants across all subgroups, limiting the statistical power of our analysis. We were also not able to analyze the association between hypothyroidism, hyperthyroidism and anemia because of the small number which restricted our analysis to TSH and FT4 levels with anemia. Such limitations must be taken into account when analyzing the results.

## CONCLUSION

The findings of this study show that higher levels of FT4 are associated with lower odds of anemia. No associations were found between TSH levels and anemia. In analysis, according to sex, no differences were found between men and women.

## AUTHORS' CONTRIBUTIONS

**ASN:** Data curation, Formal Analysis, Visualization, Writing – original draft. **MRNC:** Formal Analysis, Software, Visualization. **AGS:** Funding acquisition, Project administration, Resources. **IJMB:** Conceptualization, Methodology, Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Zhang X, Wang X, Hu H, Qu H, Xu Y, Li Q. Prevalence and trends of thyroid disease among adults, 1999-2018. Endocr Pract. 2023;29(11):875-80. https://doi.org/10.1016/j.eprac.2023.08.006
2. Olmos RD, Figueiredo RC, Aquino EM, Lotufo PA, Bensenor IM. Gender, race and socioeconomic influence on diagnosis and treatment of thyroid disorders in the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). Braz J Med Biol Res. 2015;48(8):751-8. https://doi.org/10.1590/1414-431X20154445
3. Kassebaum NJ, GBD 2013 Anemia Collaborators. The global burden of anemia. Hematol Oncol Clin North Am. 2016;30(2):247-308. https://doi.org/10.1016/j.hoc.2015.11.002
4. Helfand M, Redfern CC. Clinical guideline, part 2. Screening for thyroid disease: an update. American College of Physicians. Ann Intern Med. 1998;129(2):144-58. https://doi.org/10.7326/0003-4819-129-2-199807150-00020
5. Garmendia Madariaga A, Santos Palacios S, Guillén-Grima F, Galofré JC. The incidence and prevalence of thyroid dysfunction in Europe: a meta-analysis. J Clin Endocrinol Metab. 2014;99(3):923-31. https://doi.org/10.1210/jc.2013-2409
6. Wopereis DM, Puy RS, Heemst D, Walsh JP, Bremner A, Bakker SJL, et al. The relation between thyroid function and anemia: a pooled analysis of individual participant data. J Clin Endocrinol Metab. 2018;103(10):3658-67. https://doi.org/10.1210/jc.2018-00481
7. Vliet NA, Kamphuis AEP, Elzen WPJ, Blauw GJ, Gussekloo J, Noordam R, et al. Thyroid function and risk of anemia: a multivariable-adjusted and Mendelian randomization analysis in the UK biobank. J Clin Endocrinol Metab. 2022;107(2):e643-52. https://doi.org/10.1210/clinem/dgab674
8. Floriani C, Feller M, Aubert CE, M'Rabet-Bensalah K, Collet TH, Elzen WPJ, et al. Thyroid dysfunction and anemia: a prospective cohort study and a systematic review. Thyroid. 2018;28(5):575-82. https://doi.org/10.1089/thy.2017.0480
9. Puy RS, Poortvliet RKE, Mooijaart SP, Stott DJ, Quinn T, Sattar N, et al. No effect of levothyroxine on hemoglobin in older adults with subclinical hypothyroidism: pooled results from 2 randomized controlled trials. J Clin Endocrinol Metab. 2022;107(6):e2339-47. https://doi.org/10.1210/clinem/dgac106
10. Schmidt MI, Duncan BB, Mill JG, Lotufo PA, Chor D, Barreto SM, et al. Cohort profile: longitudinal study of adult health (ELSA-Brasil). Int J Epidemiol. 2015;44(1):68-75. https://doi.org/10.1093/ije/dyu027
11. Aquino EM, Barreto SM, Bensenor IM, Carvalho MS, Chor D, Duncan BB, et al. Brazilian Longitudinal Study of Adult Health (ELSA-Brasil): objectives and design. Am J Epidemiol. 2012;175(4):315-24. https://doi.org/10.1093/aje/kwr294
12. Bensenor IM, Griep RH, Pinto KA, Faria CP, Felisbino-Mendes M, Caetano EI, et al. [Routines of organization of clinical tests

and interviews in the ELSA-Brasil investigation center]. Rev Saude Publica. 2013;47(Suppl 2):37-47. https://doi.org/10.1590/s0034-8910.2013047003780

13. Kim M, Kim BH, Lee H, Jang MH, Kim JM, Kim EH, et al. Association between Serum free thyroxine and anemia in euthyroid adults: a nationwide study. Endocrinol Metab (Seoul). 2020;35(1):106-14. https://doi.org/10.3803/EnM.2020.35.1.106
14. Wang M, Lu X, Zheng X, Zhu X, Liu J. Associations among thyroid hormone levels and mean corpuscular volume in adults in the US: a cross-sectional examination of the NHANES 2007-2012 dataset. Medicine (Baltimore). 2024;103(10):e37350. https://doi.org/10.1097/MD.0000000000037350
15. Antoniadou C, Gavriilidis E, Ritis K, Tsilingiris D. Anemia in diabetes mellitus: pathogenetic aspects and the value of early erythropoietin therapy. Metabol Open. 2025;25:100344. https://doi.org/10.1016/j.metop.2024.100344
16. Babaker AS, Al-Qahtani RM. Frequency and characteristics of anemia in hypothyroid patients: a cross-sectional study. J Family Community Med. 2025;32(1):59-64. https://doi.org/10.4103/jfcm.jfcm_156_24
17. Silva CSS, Bensenor IM, Goulart AC, Lotufo PA, Santos IS. Anaemia and its causes at the Brazilian Longitudinal Study of Adult Health: a cross-sectional analysis of baseline data. Fam Pract. 2024;41(6):892-900. https://doi.org/10.1093/fampra/cmad102
18. Mandefro B, Kelem A, Gebreegziabher ZA, Gelaw Y, Shiferaw E, Adane T. Hematological abnormalities and associated factors among patients with thyroid hormone dysfunction at the University of Gondar Comprehensive Specialized Hospital, Northwest Ethiopia. PLoS One. 2025;20(5):e0322748. https://doi.org/10.1371/journal.pone.0322748
19. Garofalo V, Condorelli RA, Cannarella R, Aversa A, Calogero AE, Vignera S. Relationship between iron deficiency and thyroid function: a systematic review and meta-analysis. Nutrients. 2023;15(22):4790. https://doi.org/10.3390/nu15224790

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260500

# Self-efficacy in remitted bipolar disorder: associations with employment and self-esteem

Esra Yalim[1]*, Cansu Unsal Mavi[2]

## SUMMARY

**OBJECTIVE:** The aim of this study was to examine general self-efficacy in relation to perceived social support, self-esteem, and employment status among individuals with remitted bipolar disorder.

**METHODS:** This cross-sectional study included 67 patients with bipolar disorder in remission for at least eight weeks. Sociodemographic data were collected with a semi-structured form, and participants completed the Rosenberg Self-Esteem Scale, Multidimensional Scale of Perceived Social Support, and General Self-Efficacy Scale. Correlation analyses, group comparisons, and an exploratory multiple linear regression analysis were conducted.

**RESULTS:** Self-esteem was positively correlated with perceived social support ($r=0.390$, $p<0.01$) and general self-efficacy ($r=0.604$, $p<0.001$). Employed participants had higher perceived social support, self-esteem, and general self-efficacy than unemployed participants (all $p<0.05$). In an exploratory multiple linear regression model, self-esteem and employment status remained associated with general self-efficacy after adjustment for perceived social support, age, and gender.

**CONCLUSION:** Self-esteem and employment status were associated with general self-efficacy in patients with remitted bipolar disorder . These preliminary findings highlight the potential relevance of psychosocial and vocational factors, although longitudinal studies are needed to clarify the directionality of these relationships.

**KEYWORDS:** Bipolar disorder. Employment. Self efficacy. Self esteem. Social support.

## INTRODUCTION

Bipolar disorder (BD) is a chronic psychiatric disorder characterized by recurrent episodes of mania and depression that adversely affect mood, cognition, and overall functioning[1]. Notably, psychosocial functional impairments may persist even during extended periods of euthymia[2]. BD is a major contributor to global disability, with unemployment rates consistently exceeding those observed in the general population[3]. Employment status has been associated with illness management and recovery trajectories in individuals with BD[4]. Accordingly, there is increasing recognition of the importance of identifying psychosocial factors that may support functional recovery in this population.

Self-esteem is defined as an individual's subjective evaluation of self-worth[5]. Individuals with BD have been shown to exhibit lower levels of self-esteem compared with healthy controls[6]. In addition, reduced perceived social support has been associated with greater symptom severity across psychiatric conditions, including BD[7], and is also closely related to self-esteem[8]. Self-efficacy, conceptualized as the belief in one's capacity to initiate and sustain goal-directed behaviors, has been linked to quality of life in BD[9,10]. Higher self-efficacy has been associated with better psychological well-being and functional outcomes[11]. Both self-esteem and self-efficacy are well-established correlates of mental health outcomes[12].

Self-esteem, perceived social support, and self-efficacy are distinct but interrelated psychosocial constructs relevant to functional recovery in BD. Although these constructs have been examined separately, their combined associations, particularly in relation to employment status, remain insufficiently explored. Therefore, this study aimed to examine the relationships among perceived social support, self-esteem, and general self-efficacy in individuals with remitted BD and to determine whether these variables differ according to employment status.

[1]Çankırı State Hospital, Department of Psychiatry – Çankırı, Türkiye.
[2]Kartal Dr. Lütfi Kırdar City Hospital, Department of Psychiatry – Istanbul, Türkiye.
*Corresponding author: yalim.esra06@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on May 10, 2026. Accepted on July 04, 2026.
Scientific Editor: José Maria Soares Júnior

# METHODS

## Participants

A total of 67 patients diagnosed with BD according to DSM-5 criteria and considered clinically remitted were recruited from the psychiatry outpatient clinic of Çankırı State Hospital. Remission was determined by the treating psychiatrist based on clinical evaluation and was defined as the absence of acute mood symptoms and clinical stability for at least eight weeks[13]. Inclusion criteria comprised age between 18 and 65 years, a diagnosis of BD in remission, and willingness to participate in the study. Exclusion criteria included illiteracy, absence of remission, and the presence of systemic or neurological conditions that could potentially impair cognitive functioning.

## Data collection tools

Sociodemographic Data Form: A semi-structured questionnaire was used to obtain information on participants' age, gender, education level, marital status, and psychiatric history.

Rosenberg Self-Esteem Scale (RSES): Self-esteem was assessed using the Rosenberg Self-Esteem Scale developed by Rosenberg and adapted into Turkish by Çuhadaroğlu. The instrument comprises 10 items rated on a 4-point Likert scale and includes both positive and negative self-esteem subdimensions. Total scores range from 0 to 30, with higher scores indicating higher levels of self-esteem[14,15].

Multidimensional Scale of Perceived Social Support (MSPSS): Perceived social support was evaluated using the MSPSS developed by Zimet et al. and validated in Turkish by Eker et al. This 12-item self-report instrument includes three subscales: family, friends, and significant other. Items are rated on a 7-point Likert scale, with higher scores indicating greater perceived social support[16,17].

General Self-Efficacy Scale (GSES): General self-efficacy was assessed using the scale developed by Sherer et al. and validated in Turkish by Yildirim and Ilhan. The 17-item version of this self-report instrument was used in the present study. Items are rated on a 5-point Likert scale, and higher total scores indicate greater self-efficacy[18,19].

## Procedure

Ethical approval was obtained from the Çankırı Karatekin University Scientific Research Ethics Committee (14/07/2025–22). Written informed consent was obtained from all participants prior to study inclusion. The study was conducted in accordance with the principles of the Declaration of Helsinki.

## Statistical analysis

Data were analyzed using SPSS version 22.0. Descriptive statistics were reported as mean±standard deviation or n (%). Pearson correlation analyses were used to examine bivariate associations, and independent-samples t-tests were used for comparisons by employment status. An exploratory multiple linear regression analysis was conducted to examine variables associated with general self-efficacy. Education level was not entered into the final regression model because of its limited variability in the present sample and to avoid overfitting in the context of the modest sample size. Statistical significance was set at $p<0.05$. A sample size calculation using G*Power 3.1 indicated that approximately 70 participants would be required for a multiple linear regression model with five predictors, assuming a medium effect size close to Cohen's $f^2=0.20$, $\alpha=0.05$, and power=0.80.

# RESULTS

## Sample characteristics

The study sample comprised 67 patients with BD. The mean age was 32.26 years (SD=9.15). The majority of participants were female (77.6%, n=52), single or separated (67.2%, n=45), and had attained a university-level education or higher (79.1%, n=53). Slightly more than half of the participants were employed (52.2%, n=35).

## Descriptive statistics

The mean total perceived social support score was 55.09 (SD=21.98), the mean self-esteem score was 25.52 (SD=6.85), and the mean general self-efficacy score was 53.40 (SD=13.94).

## Correlation analyses

Pearson correlation analyses indicated that self-esteem was positively correlated with total perceived social support ($r=0.390$, $p<0.01$). General self-efficacy was correlated with self-esteem ($r=0.604$, $p<0.001$) and demonstrated weaker but statistically significant correlations with friend support ($r=0.278$, $p<0.05$) and significant other support ($r=0.251$, $p<0.05$), whereas no significant correlation was observed with family support or total perceived social support. Detailed correlation coefficients are presented in Table 1.

## Group comparisons by employment status

Compared with unemployed participants, employed participants reported significantly higher total perceived social support scores [MSPSS: $t(65)=-3.02$, $p=0.004$], self-esteem scores

[RSES: t(65)=-4.62, p<0.001], and general self-efficacy scores [GSES: t(65)=-4.59, p<0.001]. Detailed group comparisons are presented in Table 2.

### Exploratory multiple linear regression analysis

The exploratory regression model was significant overall [$R^2$=0.420, adjusted $R^2$=0.372, F(5,61)=8.82, p<0.001]. Self-esteem was positively associated with general self-efficacy (B=1.001, SE=0.243, p<0.001), and employed participants had higher general self-efficacy scores than unemployed participants when all variables were entered simultaneously (B=7.576, SE=3.326, p=0.026). Variance inflation factor (VIF) values ranged from 1.044 to 1.514, indicating no evidence of problematic multicollinearity. Detailed regression analysis results are presented in Table 3.

## DISCUSSION

In this study, self-esteem was positively associated with both perceived social support and general self-efficacy in patients with BD in clinical remission. Employed participants reported higher levels of these psychosocial variables

**Table 1.** Correlations between perceived social support, self-esteem, and general self-efficacy.

| | Family support | Friend support | Significant other support | MSPSS | RSES | GSES |
|---|---|---|---|---|---|---|
| Family support | – | | | | | |
| Friend support | 0.592*** | – | | | | |
| Significant other support | 0.539*** | 0.717*** | – | | | |
| MSPSS | 0.806*** | 0.897*** | 0.884*** | – | | |
| RSES | 0.237 | 0.433*** | 0.330** | 0.390** | – | |
| GSES | 0.059 | 0.278* | 0.251* | 0.233 | 0.604*** | – |

GSES: General Self-Efficacy Scale; MSPSS: Multidimensional Scale of Perceived Social Support; RSES: Rosenberg Self-Esteem Scale. *p<0.05, **p<0.01, ***p<0.001.

**Table 2.** Comparison of social support, self-esteem, and self-efficacy by employment status in bipolar disorder (n=67).

| Variable | Unemployed (mean±SD) | Employed (mean±SD) | t | df | p |
|---|---|---|---|---|---|
| MSPSS | 47.77±22.61 | 63.09±18.48 | -3.02 | 65 | 0.004 |
| RSES | 22.29±6.52 | 29.06±5.36 | -4.62 | 65 | <0.001 |
| GSES | 46.86±13.02 | 60.56±11.25 | -4.59 | 65 | <0.001 |

GSES: General Self-Efficacy Scale, MSPSS: Multidimensional Scale of Perceived Social Support, RSES: Rosenberg Self-Esteem Scale; SD: standard deviation.

**Table 3.** Exploratory multiple linear regression model (n=67).

| Variable | B | SE | LLCI | ULCI | t | p | Standardized estimate (β) |
|---|---|---|---|---|---|---|---|
| Intercept | 26.432 | 7.386 | 11.662 | 41.201 | 3.579 | p<0.001 | |
| RSES | 1.001 | 0.243 | 0.516 | 1.487 | 4.123 | p<0.001 | 0.493 |
| MSPSS | -0.029 | 0.072 | -0.173 | 0.114 | -0.408 | p=0.685 | -0.046 |
| Age | -0.029 | 0.166 | -0.360 | 0.302 | -0.177 | p=0.860 | -0.019 |
| Gender (female vs. male) | 1.596 | 3.309 | -5.021 | 8.212 | 0.482 | p=0.631 | 0.114 |
| Employment status (employed vs. unemployed) | 7.576 | 3.326 | 0.926 | 14.226 | 2.278 | p=0.026 | 0.543 |

MSPSS: Multidimensional Scale of Perceived Social Support; RSES: Rosenberg Self-Esteem Scale; β: standardized regression coefficient; CI: 95% confidence interval; LLCI: lower limit of confidence interval; ULCI: upper limit of confidence interval; SE: standard error; Gender was coded as Female=1, Male=0. Employment status was coded as employed=1, unemployed=0.

than unemployed participants. In the exploratory adjusted model, self-esteem and employment status remained associated with general self-efficacy.

Consistent with previous evidence[8,20], the present study found a positive association between perceived social support and self-esteem. These findings suggest that social support and self-esteem are interrelated psychosocial constructs in BD.

Self-efficacy has been recognized as an important correlate of mental health outcomes[21], and previous clinical research has reported positive associations among self-efficacy, self-esteem, and social support[11]. In the present study, self-esteem remained associated with general self-efficacy in the exploratory regression model. This finding is consistent with evidence that self-esteem and self-efficacy are closely related dimensions of self-perception[20] and may reflect the link between self-evaluation and perceived capability. However, it should be interpreted cautiously because the two constructs may share psychological variance.

General self-efficacy was not significantly associated with total perceived social support; only friend support and significant-other support showed small associations, whereas family support was not significant. This pattern may suggest that peer and significant-other support are more closely related to perceived competence than family support in this adult sample, although this interpretation remains tentative. Given the content of the GSES, which assesses perceived competence, initiative, persistence, and coping ability[18], these exploratory subscale findings may reflect the relevance of specific interpersonal sources of support rather than overall perceived support. However, they should not be interpreted as evidence of a robust or mediated relationship between social support and self-efficacy.

Employment has been associated with psychosocial functioning in BD[22]. In the present study, employed participants reported higher self-esteem, self-efficacy, and perceived social support than unemployed participants, and employment status remained associated with general self-efficacy in the exploratory regression model. This finding is consistent with evidence linking employment variables to self-esteem in severe mental illness[23] and with studies showing associations between self-efficacy and functional impairment in schizophrenia[24]. However, given the cross-sectional design, the absence of data on premorbid employment status and current subthreshold symptoms, and the relatively homogeneous sample, the directionality of this association remains uncertain. Thus, employment status should be regarded as a contextual correlate rather than a causal factor, and the magnitude of this association should be interpreted cautiously.

### Limitations

This study has several limitations. Its cross-sectional design precludes causal inference, and the observed associations, particularly between employment status and self-efficacy, may be bidirectional. The relatively small, single-center sample limits generalizability and may have reduced the power to detect small effects, especially in subgroup analyses; therefore, non-significant findings should be interpreted with caution. In addition, recruitment from a single center and participation on a voluntary basis may have introduced selection bias. Reliance on self-report measures may have introduced response bias. The predominance of female and highly educated participants may also have resulted in a more homogeneous and higher-functioning sample, potentially influencing regression estimates and limiting the applicability of the findings to more diverse clinical populations. Moreover, the relatively wide confidence interval for the employment-status coefficient suggests imprecision in the estimated magnitude of this association, and the corresponding regression estimate should therefore be interpreted cautiously. Several relevant confounders, including illness duration, number of previous episodes, premorbid employment history, current subthreshold mood symptoms, and psychotropic medication use, were not assessed. Remission status was determined clinically without standardized clinician-rated mood scales, which limits the objectivity of remission assessment. Finally, variables such as self-esteem and self-efficacy may fluctuate over time, which could not be captured in the present design.

## CONCLUSION

In conclusion, self-esteem was positively associated with both perceived social support and general self-efficacy in patients with clinically remitted BD. In an exploratory adjusted model, self-esteem and employment status remained associated with general self-efficacy. Given the cross-sectional design, modest sample size, and clinically determined remission status, these findings should be interpreted cautiously, and longitudinal studies with standardized mood assessments are needed to clarify directionality.

## ACKNOWLEDGEMENTS

We thank the participants of the study.

## ETHICS APPROVAL

Ethical approval was obtained from the Health Sciences Ethics Committee of Çankırı Karatekin University (14/07/2025-22).

The procedures used in this study adhere to the tenets of the Declaration of Helsinki.

## AUTHORS' CONTRIBUTIONS

**EY:** Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Writing – review & editing.

**CUM:** Formal Analysis, Supervision, Validation, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. American Psychiatric Association. DSM-5-TR ruhsal bozukluklarin tanisal ve sayimsal El kitabi. Köroğlu E, editor. 5th ed. Ankara: Esenkal Yayıncılık; 2025.
2. Bonnín CM, Montejo L, Torrent C, Sánchez-Moreno J, Diego-Adeliño J, Solé B, et al. Psychosocial functional recovery in patients with bipolar disorder in remission: which symptoms hold them back? J Affect Disord. 2026;403:121442. https://doi.org/10.1016/j.jad.2026.121442
3. Sletved KSO, Coello K, Stanislaus S, Kjærstad HL, Melbye SA, Faurholt-Jepsen M, et al. Socio-economic status and functioning in patients newly diagnosed with bipolar disorder and their unaffected siblings - results from a cross-sectional clinical study. J Affect Disord. 2022;310:404-11. https://doi.org/10.1016/j.jad.2022.05.023
4. Uzun G, Aydın Z, Kayaaslan B, Lok N. The relationship of functioning and life satisfaction with illness management and recovery in patients with bipolar disorder: a cross-sectional study. Community Ment Health J. 2024;60(6):1191-202. https://doi.org/10.1007/s10597-024-01270-x
5. Muris P, Otgaar H. Self-esteem and self-compassion: a narrative review and meta-analysis on their links to psychological problems and well-being. Psychol Res Behav Manag. 2023;16:2961-75. https://doi.org/10.2147/PRBM.S402455
6. Quinlivan E, Renneberg B, Schreiter S, Friedel E, Shmuilovich O, Stamm T. Better than expected: the gap between self-reported and objective measures of cognitive performance in remitted bipolar disorder. Front Psychiatry. 2023;14:1258303. https://doi.org/10.3389/fpsyt.2023.1258303
7. Wang J, Mann F, Lloyd-Evans B, Ma R, Johnson S. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. BMC Psychiatry. 2018;18(1):156. https://doi.org/10.1186/s12888-018-1736-5
8. Zhou X, Wu X, Zhen R. Self-esteem and hope mediate the relations between social support and post-traumatic stress disorder and growth in adolescents following the Ya'an earthquake. Anxiety Stress Coping. 2018;31(1):32-45. https://doi.org/10.1080/10615806.2017.1374376
9. Tan FCJH, Oka P, Dambha-Miller H, Tan NC. The association between self-efficacy and self-care in essential hypertension: a systematic review. BMC Fam Pract. 2021;22(1):44. https://doi.org/10.1186/s12875-021-01391-2
10. Abraham KM, Miller CJ, Birgenheir DG, Lai Z, Kilbourne AM. Self-efficacy and quality of life among people with bipolar disorder. J Nerv Ment Dis. 2014;202(8):583-8. https://doi.org/10.1097/NMD.0000000000000165
11. Rippon D, Shepherd J, Wakefield S, Lee A, Pollet TV. The role of self-efficacy and self-esteem in mediating positive associations between functional social support and psychological wellbeing in people with a mental health diagnosis. J Ment Health. 2024;33(6):721-30. https://doi.org/10.1080/09638237.2022.2069695
12. Moon JY, Kim JH. Association between self-esteem and efficacy and mental health in people with disabilities. PLoS One. 2021;16(10):e0257943. https://doi.org/10.1371/journal.pone.0257943
13. Grunze H, Vieta E, Goodwin GM, Bowden C, Licht RW, Möller HJ, et al. The World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for the biological treatment of bipolar disorders: update 2012 on the long-term treatment of bipolar disorder. World J Biol Psychiatry. 2013;14(3):154-219. https://doi.org/10.3109/15622975.2013.770551
14. Rosenberg M. Rosenberg Self-Esteem Scale (RSE). Acceptance and commitment therapy Measures package. 1965;61:18.
15. Çuhadaroğlu F. Adolesanlarda benlik saygısı. Ankara: Hacettepe Üniversitesi Tıp Fakültesi; 1986.
16. Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. J Pers Assess. 1988;52(1):30-41. https://doi.org/10.1207/s15327752jpa5201_2
17. Eker D, Arkar H, Yaldiz H. Çok boyutlu algilanan sosyal destek ölçeği'nin gözden geçirilmiş formunun faktör yapisi, geçerlik ve güvenirliği. Turk Psikiyatri Derg. 2001;12(1):17-25.
18. Sherer M, Maddux JE, Mercandante B, Prentice-Dunn S, Jacobs B, Rogers RW. The self-efficacy scale: construction and validation. Psychol Rep. 1982;51(2):663-71. https://doi.org/10.2466/pr0.1982.51.2.663
19. Yildirim F, Ilhan IO. The validity and reliability of the General Self-Efficacy Scale-Turkish form. Turk Psikiyatri Derg. 2010;21(4):301-8. PMID: 21125505
20. Ozer S. Social support, self-efficacy, self-esteem, and well-being during COVID-19 lockdown: a two-wave study of Danish students. Scand J Psychol. 2024;65(1):42-52. https://doi.org/10.1111/sjop.12952
21. Ruotolo F, Ruggiero G, Cattaneo Z, Arioli M, Candini M, Frassinetti F, et al. Psychological reactions during and after a lockdown: self-efficacy as a protective factor of mental health. Int J Environ Res Public Health. 2023;20(17):6679. https://doi.org/10.3390/ijerph20176679
22. Lomastro MJ, Valerio MP, Blasco MB, Tagni MF, Martino DJ. Predictors of high psychosocial functioning in bipolar disorder. J Nerv Ment Dis. 2020;208(11):904-7. https://doi.org/10.1097/NMD.0000000000001224
23. Deng J, Sarraf L, Hotte-Meunier A, Sauvé G. Relationship between self-esteem and employment in people with severe mental illness: a systematic review and meta-analysis. J Voc Rehab. 2023;59(3):273-83. https://doi.org/10.3233/JVR-230
24. Cardenas V, Abel S, Bowie CR, Tiznado D, Depp CA, Patterson TL, et al. When functional capacity and real-world functioning converge: the role of self-efficacy. Schizophr Bull. 2013;39(4):908-16. https://doi.org/10.1093/schbul/sbs004

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260564

# A comparative analysis of semiannual versus annual rituximab maintenance in patients with refractory idiopathic inflammatory myopathies: a retrospective longitudinal real-world study

Rodrigo Lorenzetti Serrano[1], Fernando Henrique Carlos de Souza[1], Renata Miossi[1], Marlise Sítima Mendes Simões Faria[1], Samuel Katsuyuki Shinjo[1]*

## SUMMARY

**OBJECTIVE:** The aim of this study was to explore the real-world effectiveness of semiannual versus annual rituximab maintenance administration in patients with refractory idiopathic inflammatory myopathies.

**METHODS:** In this single-center, retrospective longitudinal study (2010–2025), 45 adults with refractory idiopathic inflammatory myopathies received two initial rituximab infusions (1 g on days 0 and 15). Maintenance rituximab was administered every 6 months (semiannual group, n=14) or every 12 months (annual group, n=31). Baseline immunosuppressive therapy was maintained, and glucocorticoid tapering was guided by clinical and laboratory parameters. Outcomes at months 6–12 and 18–24 included the need for rescue therapies: intravenous methylprednisolone (3 g total) and immunoglobulin (2 g/kg), or prednisone dose escalation.

**RESULTS:** Baseline demographics, refractory idiopathic inflammatory myopathies subtypes, rituximab indications (cutaneous, muscular, articular, and/or pulmonary activity), and peri- rituximab medications (within 2 months before/after infusion) did not differ significantly between groups ($p>0.05$). No statistically significant differences in the frequency of rescue therapy use at 6–12 and 18–24 months were detected between the two regimens. Intragroup analysis demonstrated significant reductions in relapse rates and rescue therapy requirements over time ($p<0.05$), suggesting improvement in both cohorts.

**CONCLUSION:** In this exploratory cohort, no statistically significant differences were detected between semiannual and annual rituximab maintenance regimens with respect to clinical outcomes, although numerically higher relapse rates were observed in the annual group. These findings suggest that extended rituximab intervals may be feasible in resource-limited settings, but further prospective studies are needed to confirm these observations.

**KEYWORDS:** Myositis. Idiopathic inflammatory myopathies. Rituximab. Treatment.

## INTRODUCTION

Idiopathic inflammatory myopathies (IIMs), also referred to as systemic autoimmune myopathies, comprise a heterogeneous group of disorders primarily affecting skeletal muscle, though extra-muscular involvement—such as cutaneous, articular, and pulmonary manifestations—is also common[1]. Management typically involves glucocorticoids, intravenous immunoglobulin (IVIG), and various immunosuppressive or immunomodulatory agents, administered either as monotherapy or in combination[2]. In severe or treatment-refractory cases, the use of biologic therapies or small-molecule agents is recommended[3,4].

Among the biologic agents employed in IIM, rituximab (RTX) has emerged as a key therapeutic option for refractory disease, with favorable clinical outcomes reported across several studies[5-7]. The conventional RTX regimen for IIM involves two 1 g infusions on days 0 and 15, repeated every 6 months, though an alternative lymphoma-based protocol (375 mg/m$^2$ weekly for 4 weeks) is also used. Recent studies have explored alternative dosing schedules: Janardana et al.[8] demonstrated comparable efficacy using 500 mg on days 0 and 15 every 6 months over 12 months of follow-up, whereas Gamba et al.[9] reported similar outcomes with a single 1 g dose every 6 months during longer-term observation. Ahn et al.[10] observed a mean response rate of 68.8% over 6 months, despite heterogeneous dosing intervals.

These findings are particularly relevant in public healthcare contexts, where limited availability may preclude adherence to standard dosing intervals. Moreover, the long-term impact of extended RTX dosing schedules remains insufficiently studied.

[1]Universidade de São Paulo, Faculdade de Medicina FMUSP, Division of Rheumatology – São Paulo (SP), Brazil.
*Corresponding author: samuel.shinjo@usp.br

Conflicts of interest: the authors declare there is no conflicts of interest. Funding: this work was funded by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq #303379/2018-9 to SKS) and Faculdade de Medicina da USP to SKS.
Received on June 01, 2026. Accepted on July 13, 2026.
Scientific Editor: Roseli Nomura

Accordingly, this study aimed to evaluate disease activity and relapse rates over a 24-month period in patients receiving standard RTX doses either semiannually or annually in a real-world setting.

## PATIENTS AND METHODS

This single-center, retrospective longitudinal study was conducted between 2010 and 2025 and included adult patients diagnosed with refractory idiopathic inflammatory myopathies (rIIM) as follows: dermatomyositis (DM), according to the 2017 EULAR/ACR criteria[11]; clinically amyopathic dermatomyositis (CADM), based on Gerami et al.'s definition[12]; anti-synthetase syndrome (ASyS), according to Connors et al.[13]; and immune-mediated necrotizing myopathy (IMNM), following the criteria by Allenbach et al[14].

Refractory IIM was defined by inadequate response to glucocorticoids and at least two immunosuppressive agents (e.g., azathioprine, methotrexate, mycophenolate mofetil, cyclosporine, or leflunomide), administered sequentially or concomitantly for ≥3 months[15].

The study protocol was approved by the local ethics committee, and all participants provided written informed consent.

As exclusion criteria, patients with inclusion body myositis or other non-IIM myopathies, those with severe infections requiring treatment discontinuation, and individuals receiving RTX doses or intervals inconsistent with the study protocol were excluded.

Data collection. Patient data were extracted from electronic medical records using standardized forms, capturing the following variables: demographic: age, sex, ethnicity; clinical: disease duration, muscle strength, dyspnea, arthritis, cutaneous involvement; laboratory data: C-reactive protein (CRP), erythrocyte sedimentation rate (ESR), complete blood count (CBC), creatinine, transaminases, creatine phosphokinase (CK); imaging: chest computed tomography (CT), pulmonary function tests, muscle magnetic resonance imaging (MRI), and joint ultrasound (when indicated); treatment history: previous and current medications, duration of use, and adverse events.

Disease activity was evaluated across four domains:

- Muscular: clinical worsening on Manual Muscle Testing (MMT-8)[16] or CK elevation above baseline or ≥ three upper limits of normal;
- Pulmonary: new or worsening dyspnea, decline in pulmonary function tests or imaging, excluding infections;
- Cutaneous: physical examination findings consistent with active disease[17];
- Articular: presence of arthralgia or arthritis, confirmed with imaging if necessary.

Disease flare/relapse was assessed according to the International Myositis Assessment & Clinical Studies Group (IMACS) concept of clinical worsening after a period of disease stability or improvement, involving one or more core disease activity domains[18]. In this study, relapse was pragmatically operationalized as the need for rescue therapy during ongoing RTX treatment, including methylprednisolone pulse therapy (1 g/day for 3 days), intravenous human immunoglobulin (2 g/kg), and/or initiation or escalation of prednisone dosage. To improve consistency, rescue interventions were considered relapse-related only when accompanied by documented clinical and/or laboratory evidence of worsening disease in at least one predefined organ domain (muscular, pulmonary, cutaneous, or articular), after exclusion of alternative causes such as infection or drug toxicity. Nevertheless, this outcome represents real-world physician-driven therapeutic decisions and should be interpreted as a pragmatic clinical marker rather than a validated composite measure of global disease activity.

Timepoints analyzed. Two primary periods were analyzed: peri-RTX (the 2 months before and after initial RTX infusion), and post-RTX (subdivided into months 6–12 and 18–24).

At each timepoint, relapse events and the affected clinical domains were recorded.

RTX regimens. All patients received two 1 g infusions of RTX on days 0 and 15: semiannual group (received repeated doses every 6 months), and annual group (received repeated doses every 12 months, due to limited drug availability).

RTX infusions were withheld in cases of previous adverse reactions, active infections, or hypogammaglobulinemia. Immunosuppressive therapy remained unchanged post-RTX, and glucocorticoids were tapered based on clinical response.

Statistical analysis. Normality was assessed using the Shapiro-Wilk test. Continuous variables are expressed as mean±standard deviation (SD) or median [interquartile range (IQR): 25th–75th], as appropriate; categorical variables are presented as frequency (%). Intergroup comparisons used Student's t-test or Mann-Whitney U test for continuous variables, and chi-square or Fisher's exact test for categorical data. Proportions were analyzed with 95%CI using a binomial distribution. Statistical significance was set at $p<0.05$.

## RESULTS

Out of an initial cohort of 84 patients, one was excluded due to a severe infection after the first dose, one for receiving a non-standard RTX dose (e.g., other than 1 g on days 0 and 15), six for having received RTX for less than 24 months, and 31 for being treated with administration intervals differing from

the semiannual or annual protocols (Figure 1). Consequently, 45 adult patients with rIIM were included in the final analysis: three with CADM, nine with IMNM, 11 with DM, and 22 with ASyS. Among these, 31 patients received RTX annually and 14 semiannually (Table 1). We performed an aggregated, hypothesis-generating analysis. Of the 45 patients, none had hypogammaglobulinemia, infections, hospitalization, infusion reactions, discontinuation, or mortality.

The mean age was similar between groups: 44.9 years in the annual RTX group and 46.7 years in the semiannual group (p=0.540). The distribution of rIIM subtypes did not differ significantly between groups: CADM (6.5 vs. 7.1%), IMNM (22.6 vs. 14.3%), DM (22.6 vs. 28.6%), and ASyS (48.4 vs. 50.0%)—all with p>0.05.

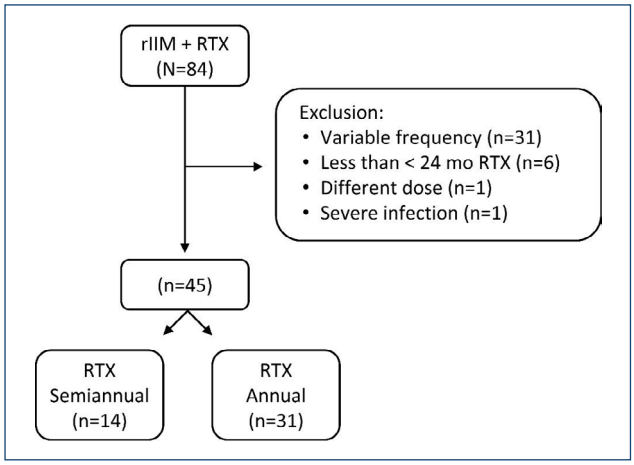



**Figure 1.** Flowchart of patient selection for the study cohort.

**Table 1.** Demographic and disease characteristics.

| | Annual RTX (n=31) | Semiannual RTX (n=14) | p-value | Effect size/odds ratio (95%CI) |
|---|---|---|---|---|
| Mean age (SD), years | 44.6 (±10.8) | 46.3 (±8.4) | 0.531[a] | d=0.120 (-0.243–0.453) |
| Sex (female, %) | 23 (74.0) | 9 (64.0) | 0.497[b] | OR 0.626 (0.161–2.433) |
| Median disease duration (IQR 25th-75th), years | 3 (2.0–5.5) | 3 (2.0–5.8) | 0.872[c] | $r_{rb}$=-0.032 (-0.381–0.324) |
| IIM subtype (%) | | | | |
| DM | 7 (22.6) | 4 (28.6) | 0.717[d] | OR 0.73 (0.14–3.48) |
| CADM | 2 (6.5) | 1 (7.1) | >0.999[d] | OR 0.90 (0.01–15.70) |
| AsyS | 15 (48.4) | 7 (50.0) | >0.999[d] | OR 0.94 (0.23–3.78) |
| IMNM | 7 (22.6) | 2 (14.3) | 0.698[d] | OR 1.75 (0.26–18.04) |
| Antibody (%) | | | | |
| Anti-Mi2 | 1 (3.0) | 1 (7.1) | | |
| Anti-SAE | 1 (3.0) | 0 | | |
| Anti-Ro52 | 2 (6.0) | 2 (14.2) | | |
| Anti-NXP2 | 1 (3.0) | 0 | | |
| Anti-TIF1g | 1 (3.0) | 0 | | |
| Anti-MDA5 | 0 | 1 (7.1) | | |
| Anti-Jo1 | 8 (25.8) | 6 (42.8) | | |
| Anti-PL7/PL12 | 5 (16.0) | 1 (7.1) | | |
| Anti-SRP | 3 (9.6) | 0 | | |
| Anti-HMG-CoA | 1 (3.0) | 0 | | |
| Anti-PM/Scl | 2 (6.0) | 0 | | |
| Peri-rituximab (%) | | | | |
| Methylprednisolone pulse therapy | 7 (22.6) | 3 (21.4) | >0.999[d] | OR 1.07 (0.19–0.34) |
| Intravenous immunoglobulin | 4 (12.9) | 4 (28.6) | 0.231[d] | OR 0.37 (0.05–2.19) |
| Prednisone use | 26 (83.9) | 14 (100) | 0.305[d] | OR 0.54 (0.03–5.23) |
| Median baseline prednisone dose (IQR 25th-75th) mg | 7.5 (5.0–12.5) | 12.5 (7.5–20.0) | 0.076[c] | $r_{rb}$=0.338 (-0.028–0.624) |
| Concomitant immunosuppressor (%) | | | | |
| 0–1 | 7 (22.5) | 4 (28.6) | 0.720[d] | OR 1.37 (0.34–5.50) |
| ≥2 | 24 (77.4) | 10 (71.4) | | |

ASyS: anti-synthetase syndrome; CADM: clinically amyopathic dermatomyositis; DM: dermatomyositis; IIM: idiopathic inflammatory myopathies; IVIG: intravenous human immunoglobulin; IMNM: immune-mediated necrotizing myopathy; RTX: rituximab; SD: standard deviation; IQR: Interquartile range; CI: confidence interval. [a]Student's t-test for independent samples. [b]Chi-Square test of independence. [c]Mann-Whitney U test. [d]Fisher's exact test. Estimates calculated using Haldane-Anscombe correction due to zero cell counts.

The predominant clinical domains prompting RTX initiation—muscular, cutaneous, articular, or pulmonary involvement—were similarly distributed between groups (p>0.05). Regarding median CK levels, there were no differences between groups pre-RTX (167 U/L [137–753] vs. 191.5 U/L [107–1168] for annual and semiannual, respectively; p=0.741), and post-RTX (182 U/L [65–345] vs. 157 [83–260], p=0.750). We also did not find a difference in median MMT-8 score between groups pre-RTX (80 [60–80] vs. 76.5 [69–80], p=0.596), and post-RTX (80 [73–80] vs. 80 [78.5–80], p=0.168).

Peri-RTX treatments also did not differ significantly: methylprednisolone pulses (22.6 vs. 21.4% in the annual and semiannual group, respectively; p>0.999), IVIG (12.9 vs. 28.6%; p=0.231), and prednisone use (83.9 vs. 100%; p=0.305).

During the 6–12-month follow-up period, the need for rescue therapies remained comparable between groups. In the annual RTX group, 22.6% required methylprednisolone, 6.5% received IVIG, and 16.1% required prednisone escalation, compared to 7.1% for each intervention in the semiannual group (p>0.05 for all comparisons). Regarding methylprednisolone, the proportion of patients was numerically higher in the annual group, although the difference was not statistically significant. Muscular relapses occurred in 25.8% of patients in the annual group compared to 7.1% in the semiannual group. Although this numerical difference favored the semiannual regimen, it did not reach statistical significance (p=0.236) (Table 2).

Between months 18 and 24, a global reduction in rescue interventions was observed. In the annual group, 9.7% received methylprednisolone, 6.5% IVIG, and 3.2% prednisone. In the semiannual group, no patients required methylprednisolone or prednisone, and only 7.1% received IVIG (p>0.05). Relapses were limited to isolated manifestations, with no statistically significant differences.

Intragroup analysis demonstrated a significant reduction in the frequency of rescue therapy use between the 6–12 and 18–24-month follow-up periods within each treatment group, demonstrating a significant reduction over time (p<0.05 for both groups), indicating comparable efficacy and sustained clinical improvement regardless of RTX administration frequency (Table 2).

**Table 2.** Rescue therapies and relapse rates at 6–12 and 18–24 months.

| | Annual RTX (n=31) | Semiannual RTX (n=14) | p-value | Effect size/odds ratio (95%CI) |
|---|---|---|---|---|
| Post-rituximab (%) | | | | |
| Period: Months 6–12 of follow-up | | | | |
| Methylprednisolone pulse therapy | 7 (22.6) | 1 (7.1) | 0.402[d] | OR 3.79 (0.42–34.27) |
| IVIG | 2 (6.5) | 1 (7.1) | >0.999[d] | OR 0.90 (0.07–10.79) |
| Prednisone | 5 (16.1) | 1 (7.1) | 0.648[d] | OR 2.50 (0.26–23.67) |
| Disease activity (relapse) | | | | |
| Cutaneous | 2 (6.5) | 1 (7.1) | >0.999[d] | OR 0.90 (0.07–10.79) |
| Muscular | 8 (25.8) | 1 (7.1) | 0.236[d] | OR 4.52 (0.51–40.29) |
| Articular | 2 (6.5) | 0 | >0.999[d] | OR 2.46* (0.11–54.60) |
| Pulmonary | 1 (3.2) | 0 | >0.999[d] | OR 1.43* (0.05–37.20) |
| Period: Months 18–24 of follow-up | | | | |
| Methylprednisolone pulse therapy | 3 (9.7) | 0 | >0.999[d] | OR 3.56* (0.17–73.71) |
| IVIG | 2 (6.5) | 1 (7.1) | 0.488[d] | OR 0.90 (0.07–10.79) |
| Prednisone | 1 (3.2) | 0 | >0.999[d] | OR 1.43* (0.05–37.20) |
| Disease activity (relapse) | | | | |
| Cutaneous | 0 | 0 | >0.999[d] | OR 0.46* (0.01–24.37) |
| Muscular | 2 (6.5) | 0 | >0.999[d] | OR 2.46* (0.11–54.60) |
| Articular | 1 (3.2) | 1 (7.1) | 0.262[d] | OR 0.43 (0.03–7.47) |
| Pulmonary | 1 (3.2) | 1 (7.1) | 0.262[d] | OR 0.43 (0.03–7.47) |

IVIG: intravenous human immunoglobulin; RTX: rituximab; CI: confidence interval. [d]Fisher's exact test. *Estimates calculated using Haldane-Anscombe correction due to zero cell counts.

## DISCUSSION

This study evaluated patients with rIIM treated with RTX administered either semiannually or annually. Over a 24-month follow-up period, no statistically significant differences were observed between the two regimens in relapse rates or the need for rescue therapies. These exploratory findings indicate that extending the RTX dosing interval to 12 months may be a feasible option for selected patients, although this approach is not currently incorporated into standard treatment protocols.

To date, most studies investigating RTX in rIIM have adopted regimens extrapolated from rheumatoid arthritis (1 g on days 0 and 15 every 6 months) or lymphoma (375 mg/m$^2$ weekly for 4 weeks, repeated semiannually). Randomized trials such as the RIM Trial, together with multiple observational studies, have supported the efficacy of semiannual RTX administration in reducing disease activity and facilitating corticosteroid tapering[8,9,19]. Standard dosing strategies have been evaluated in cohorts with DM, ASyS, and IMNM[20-23]. However, to our knowledge, no prior study has directly compared extended RTX intervals, such as annual dosing, in this patient population.

Despite the retrospective design, the use of standardized data collection instruments helped minimize information bias. The results provide preliminary evidence that annual RTX administration did not result in statistically significant increases in relapse rates or reliance on adjunctive immunomodulatory therapies. Our study used a pragmatic, treatment-based definition of relapse, which, despite being standardized through objective worsening in predefined disease domains, may carry inherent center-specific prescribing biases.

Temporal analysis revealed a general decline in the need for rescue interventions in both groups, suggesting that sustained B-cell depletion may contribute to long-term disease control regardless of dosing interval. This finding may be particularly relevant in public healthcare systems where logistical and financial constraints can limit access to biologic therapies.

It is important to emphasize that treatment individualization remains essential. Decisions regarding RTX dosing intervals should consider disease subtype, organ involvement, relapse history, and infection risk. The lack of statistically significant differences observed in this study may be attributable to the limited sample size, clinical heterogeneity, and concurrent immunosuppressive therapies. Additionally, the absence of significant changes in creatine kinase levels or MMT-8 scores in either group should be interpreted cautiously, as many patients had subtypes with minimal muscle involvement and received RTX for non-muscular manifestations.

Key limitations include the observational, non-randomized design, small sample size, absence of quality-of-life data, and lack of a validated extramuscular disease activity tool. RTX interval allocation was influenced by drug availability, and baseline treatment patterns suggest that patients in the semiannual group may have had more severe or refractory disease, potentially confounding group comparisons. This factor should be considered when interpreting the numerically lower relapse rates in the semiannual group.

Prospective, controlled trials with larger cohorts are needed to further evaluate the long-term safety and efficacy of RTX across different dosing schedules.

## CONCLUSION

In this real-world study, no statistically significant differences were detected between semiannual and annual RTX regimens in patients with rIIM over 24 months, with respect to relapse rates or need for rescue therapies. While numerically higher relapse rates were observed in the annual group, these did not reach statistical significance. These exploratory findings indicate that annual RTX administration may be a feasible option for selected patients—especially in resource-limited settings—when appropriate clinical monitoring is in place. However, larger prospective randomized studies are needed to confirm these observations and to determine the optimal RTX dosing intervals.

## ETHICS APPROVAL AND WRITTEN INFORMED CONSENT STATEMENTS

The study was approved by the Research Ethics Committee of Faculdade de Medicina da Universidade de São Paulo (protocol number 0324/2011, approved on 07/29/2011). All participants provided written informed consent prior to inclusion. The study was conducted in accordance with the Declaration of Helsinki.

## AUTHORS' CONTRIBUTIONS

**RLS:** Conceptualization, Data curation, Funding acquisition, Investigation, Writing – original draft, Writing – review & editing. **FHCS:** Conceptualization, Funding acquisition, Investigation, Writing – original draft, Writing – review & editing. **RM:** Conceptualization, Funding acquisition, Investigation, Writing – original draft, Writing – review & editing. **MSMSF:** Conceptualization, Funding acquisition, Investigation, Writing – original

draft, Writing – review & editing. **SKS:** Conceptualization, Data curation, Funding acquisition, Investigation, Supervision, Writing – original draft, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Lundberg IE, Fujimoto M, Vencovsky J, Aggarwal R, Holmqvist M, Christopher-Stine L, et al. Idiopathic inflammatory myopathies. Nat Rev Dis Primers. 2021;7(1):86. https://doi.org/10.1038/s41572-021-00321-x
2. Paik JJ, Werth VP, Chinoy H, Masri KR, Jambekar A, Hasan F, et al. Treatment guidelines for idiopathic inflammatory myopathies in adults: a comparative review. Rheumatology (Oxford). 2025;64(6):3288-302. https://doi.org/10.1093/rheumatology/keaf116
3. Campanilho-Marques R, Fonseca JE, Machado PM. Treatment of idiopathic inflammatory myopathies. Joint Bone Spine. 2025;92(6):105932. https://doi.org/10.1016/j.jbspin.2025.105932
4. Oldroyd AGS, Lilleker JB, Amin T, Aragon O, Bechman K, Cuthbert V, et al. British Society for Rheumatology guideline on management of paediatric, adolescent and adult patients with idiopathic inflammatory myopathy. Rheumatology (Oxford). 2022;61(5):1760-8. https://doi.org/10.1093/rheumatology/keac115
5. Stano S, Cacciapaglia F, Coladonato L, Lopalco G, Iannone F, Fornaro M. Real-life effectiveness of rituximab in different subsets of idiopathic inflammatory myopathies. Reumatismo. 2025;77(3). https://doi.org/10.4081/reumatismo.2025.1817
6. Rojas LO, Ramsubeik K, Sanchez-Ramos L, Motilal S, Singh JA, Kaeley GS. Rituximab treatment in adult patients with idiopathic inflammatory myositis: a systematic review and meta-analysis. J Clin Rheumatol. 2025;31(1):33-9. https://doi.org/10.1097/RHU.0000000000002151
7. Ramos-Casals M, García-Hernández FJ, Ramón E, Callejas JL, Martínez-Berriotxoa A, Pallarés L, et al. Off-label use of rituximab in 196 patients with severe, refractory systemic autoimmune diseases. Clin Exp Rheumatol. 2010;28(4):468-76. PMID: 20525449
8. Janardana R, Amin SN, Rajasekhar L, Pinto B, Kodishala C, Selvam S, et al. Low-dose rituximab is efficacious in refractory idiopathic inflammatory myopathies. Rheumatology (Oxford). 2023;62(3):1243-7. https://doi.org/10.1093/rheumatology/keac438
9. Gamba A, Depascale R, Zanatta E, Ienna L, Cruciani C, Gatto M, et al. Effectiveness and safety of low dose Rituximab as remission-maintenance treatment for patients with refractory idiopathic inflammatory myopathies: results of a retrospective study from a monocentric cohort. Clin Rheumatol. 2024;43(10):3167-74. https://doi.org/10.1007/s10067-024-07079-z
10. Ahn GY, Suh CH, Kim YG, Park YB, Shim SC, Lee SH, et al. Efficacy and safety of rituximab in Korean patients with refractory inflammatory myopathies. J Korean Med Sci. 2020;35(38):e335. https://doi.org/10.3346/jkms.2020.35.e335
11. Lundberg IE, Tjärnlund A, Bottai M, Werth VP, Pilkington C, Visser M, et al. 2017 European League Against Rheumatism/American College of Rheumatology classification criteria for adult and juvenile idiopathic inflammatory myopathies and their major subgroups. Ann Rheum Dis. 2017;76(12):1955-64. https://doi.org/10.1136/annrheumdis-2017-211468
12. Gerami P, Schope JM, McDonald L, Walling HW, Sontheimer RD. A systematic review of adult-onset clinically amyopathic dermatomyositis (dermatomyositis sine myositis): a missing link within the spectrum of the idiopathic inflammatory myopathies. J Am Acad Dermatol. 2006;54:597-613.
13. Connors GR, Christopher-Stine L, Oddis CV, Danoff SK. Interstitial lung disease associated with the idiopathic inflammatory myopathies: what progress has been made in the past 35 years? Chest. 2010;138(6):1464-74. https://doi.org/10.1378/chest.10-0180
14. Allenbach Y, Mammen AL, Benveniste O, Stenzel W, Immune-Mediated Necrotizing Myopathies Working Group. 224th ENMC International Workshop::clinico-sero-pathological classification of immune-mediated necrotizing myopathies Zandvoort, The Netherlands, 14-16 October 2016. Neuromuscul Disord. 2018;28(1):87-99. https://doi.org/10.1016/j.nmd.2017.09.016
15. Souza FHC, Miossi R, Moraes JCB, Bonfá E, Shinjo SK. Favorable rituximab response in patients with refractory idiopathic inflammatory myopathies. Adv Rheumatol. 2018;58(1):31. https://doi.org/10.1186/s42358-018-0030-z
16. Rider LG, Koziol D, Giannini EH, Jain MS, Smith MR, Whitney-Mahoney K, et al. Validation of manual muscle testing and a subset of eight muscles for adult and juvenile idiopathic inflammatory myopathies. Arthritis Care Res (Hoboken). 2010;62(4):465-72. https://doi.org/10.1002/acr.20035
17. Anyanwu CO, Fiorentino DF, Chung L, Dzuong C, Wang Y, Okawa J, et al. Validation of the cutaneous dermatomyositis disease area and severity index: characterizing disease severity and assessing responsiveness to clinical change. Br J Dermatol. 2015;173(4):969-74. https://doi.org/10.1111/bjd.13915
18. Oddis CV, Rider LG, Reed AM, Ruperto N, Brunner HI, Koneru B, et al. International consensus guidelines for trials of therapies in the idiopathic inflammatory myopathies. Arthritis Rheum. 2005;52(9):2607-15. https://doi.org/10.1002/art.21291
19. Oddis CV, Reed AM, Aggarwal R, Rider LG, Ascherman DP, Levesque MC, et al. Rituximab in the treatment of refractory adult and juvenile dermatomyositis and adult polymyositis: a randomized, placebo-phase trial. Arthritis Rheum. 2013;65(2):314-24. https://doi.org/10.1002/art.37754
20. He C, Li W, Xie Q, Yin G. Rituximab in the treatment of interstitial lung diseases related to anti-melanoma differentiation-associated gene 5 dermatomyositis: a systematic review. Front Immunol. 2022;12:820163. https://doi.org/10.3389/fimmu.2021.820163
21. Narváez J, Cañadillas E, Castellví I, Alegre JJ, Vicens-Zygmunt V, Bermudo G, et al. Rituximab in the treatment of progressive interstitial lung disease associated with the antisynthetase syndrome. Arthritis Res Ther. 2024;26(1):122. https://doi.org/10.1186/s13075-024-03353-2
22. Suh J, Amato AA. Management of immune-mediated necrotizing myopathy. Muscle Nerve. 2024;70(2):166-72. https://doi.org/10.1002/mus.28114
23. Landon-Cardinal O, Allenbach Y, Soulages A, Rigolet A, Hervier B, Champtiaux N, et al. Rituximab in the treatment of refractory anti-HMGCR immune-mediated necrotizing myopathy. J Rheumatol. 2019;46(6):623-7. https://doi.org/10.3899/jrheum.171495

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260575

# Inpatient rehabilitation outcomes in acquired brain injury: a retrospective study

Sefa Gümrük Aslan[1]*, Şehri Bariş Saklica[1], Sinem Uyar Köylü[1], Kurtuluş Köklü[1]

## SUMMARY

**OBJECTIVE:** The objective of this study was to describe the demographic, clinical, cognitive, and functional characteristics of acquired brain injury patients, analyze complications, and evaluate inpatient rehabilitation outcomes.

**METHODS:** Retrospective review of 50 acquired brain injury patients undergoing multidisciplinary inpatient rehabilitation (2019–2024). Glasgow Coma Scale, post-traumatic amnesia, cognitive scales (Mini-Mental State Examination, Montreal Cognitive Assessment, Rancho Los Amigos), Functional Ambulation Scale, and complications were assessed.

**RESULTS:** Traumatic etiology comprised 76% of cases; 72% were severe (Glasgow Coma Scale 3–8). Prolonged post-traumatic amnesia (>24 h) was significantly associated with severe injury ($p<0.001$). Spasticity (64%), contractures (42%), speech disorders (42%), and dysphagia (34%) were most prevalent. Functional Ambulation Scale score 0 decreased from 58 to 42% post-rehabilitation; Functional Ambulation Scale 3–5 increased from 24 to 46% (Wilcoxon signed-rank test, $p<0.001$).

**CONCLUSION:** Structured multidisciplinary rehabilitation was associated with measurable functional gains in acquired brain injury patients despite severe deficits and frequent complications; however, the absence of a control group precludes causal conclusions.

**KEYWORDS:** Brain injuries. Brain injuries, traumatic. Neurological rehabilitation. Treatment outcome. Recovery of function. Glasgow Coma Scale.

## INTRODUCTION

Acquired brain injury (ABI) encompasses damage to the central nervous system from both traumatic and non-traumatic etiologies, resulting in physical, cognitive, and psychosocial impairments. Traumatic brain injury (TBI) is a leading global health burden with an annual incidence of 150–300 per 100,000; males are affected two to three times more frequently, particularly those aged 18–25 years[1-5]. Non-traumatic ABI includes focal lesions, anoxic injury, tumors, aneurysms, and CNS infections[6]. ABI patients face a high risk of secondary complications—spasticity, contractures, post-traumatic epilepsy, dysphagia, and heterotopic ossification—that impair recovery[7-9]. Early, multidisciplinary rehabilitation reduces these complications, enhances neuroplasticity, and improves long-term outcomes[10-13]. This study aimed to describe the demographic, clinical, cognitive, and functional profile of ABI patients admitted to a specialized inpatient rehabilitation center and evaluate outcomes stratified by injury severity.

## METHODS

### Study design and participants

This retrospective, single-center study reviewed records of 350 ABI patients hospitalized for inpatient rehabilitation between 2019 and 2024. Inclusion criteria were as follows: age ≥18 years, confirmed ABI diagnosis, absence of comorbidities affecting cognitive or functional status, and complete medical records. Of the 350 patients screened, 300 were excluded: 128 due to missing admission and/or discharge assessment data (e.g., baseline or discharge GCS, FAS, or cognitive scale scores not recorded), 130 due to other incomplete chart/file data required for this analysis, and 42 because the patient was discharged prematurely for an acute medical emergency before completing the rehabilitation program and therefore lacked discharge outcome data. After screening, 50 patients with complete admission-to-discharge records were included. The study was approved by the Ethics Committee of Gaziler

[1]University of Health Sciences, Ankara Gaziler Physical Medicine and Rehabilitation Health Application and Research Center, Physical Medicine and Rehabilitation – Ankara, Türkiye.
*Corresponding author: sefagumrukk@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 08, 2026. Accepted on July 19, 2026.
Scientific Editor: José Maria Soares Júnior

Physical Medicine and Rehabilitation Training and Research Hospital (Date: 26.10.2023; Protocol No: 29/008).

### Data collection

Demographic variables, date of injury, initial GCS score, PTA duration, and injury etiology were recorded. PTA duration was categorized as none, <1 h, 1–24 h, or >24 h based on retrospective chart documentation; as intermediate durations were rarely specified in the medical records, all patients fell into either the none or >24-h categories. Muscle strength was evaluated using the MRC scale. Cognitive status was assessed with Mini-Mental State Examination (MMSE), Montreal Cognitive Assessment (MoCA), and the Rancho Los Amigos Scale. Functional mobility was evaluated with FAS.

### Rehabilitation protocol

All patients received a standardized neurological rehabilitation program, delivered 5 days per week for 45 min per session, regardless of severity group. The program consisted of neurophysiological exercises, balance exercises, postural training exercises, proprioceptive exercises, and ambulation and transfer exercises, in addition to ROM exercises, verticalization training, speech-cognitive therapy, ADL exercises, and hand therapy, provided under the supervision of physiatrists, physiotherapists, occupational therapists, and speech-language pathologists.

### Statistical analysis

Statistical analyses were performed using SPSS v27.0. Categorical variables were expressed as frequencies and percentages; continuous variables as means±SD. Chi-square tests were applied for categorical comparisons. Student's t-test or Mann-Whitney U test was used for continuous variables; one-way ANOVA for three-group comparisons. Normality was assessed using the Kolmogorov-Smirnov test. $p<0.05$ was considered statistically significant. Given the disparate subgroup sizes (severe n=36, moderate n=8, mild n=6), post-hoc power analysis was performed for the one-way ANOVA comparisons of cognitive outcomes using the observed effect sizes, group sizes, and an alpha of 0.05. Kolmogorov-Smirnov/Shapiro-Wilk testing indicated that MMSE and Rancho Los Amigos scores were significantly non-normally distributed in the severe subgroup (and, for Rancho Los Amigos, also in the mild subgroup), which is noted as a caveat on the corresponding ANOVA results given the small and unequal subgroup sizes. Within-group pre-post change in FAS score (admission vs. discharge) was assessed using the Wilcoxon signed-rank test given its ordinal, paired nature. An independent-samples t-test was used to compare traumatic and non-traumatic etiology groups on muscle strength and cognitive outcomes, and chi-square tests were used to compare these groups on complication frequency and FAS scores. The study was reported in accordance with the STROBE statement.

## RESULTS

### Demographic characteristics

Fifty patients were included (mean age 34.68±12.88 years; 80% male). Mean BMI was 25.22±3.83 kg/m$^2$. High school graduates comprised 62%; 52% were single.

### Clinical characteristics

Mean time since injury was 15.62±14.08 months. Traumatic etiology accounted for 76% of cases. Mean GCS was 6.5±4.17; 72% were classified as severe (GCS 3–8), 16% moderate, and 12% mild. Chi-square analysis confirmed a significant association between severity and prolonged PTA ($p<0.001$): 83.3% of severe TBI patients experienced PTA >24 h vs. 25% in the moderate and 0% in the mild group (Table 1). Mean MMSE was 23.2±4.60, MoCA 19.22±5.53, and Rancho Los Amigos 6.94±1.67.

**Table 1.** Clinical characteristics of patients with acquired brain injury (n=50).

| Variable | Value (n=50) |
|---|---|
| Time since injury (months) | 15.62±14.08 |
| Etiology, n (%) | |
| Traumatic | 38 (76) |
| Non-traumatic | 12 (24) |
| Glasgow Coma Scale (GCS) | 6.5±4.17 |
| GCS classification, n (%) | |
| Severe (3–8) | 36 (72) |
| Moderate (9–12) | 8 (16) |
| Mild (13–15) | 6 (12) |
| Post-traumatic amnesia, n (%) | |
| None | 18 (36) |
| <1 h | 0 |
| 1–24 h | 0 |
| >24 h | 32 (64) |
| Upper extremity muscle strength | 3.46±1.21 |
| Lower extremity muscle strength | 3.42±1.07 |
| MMSE | 23.2±4.60 |
| MoCA | 19.22±5.53 |
| Rancho Los Amigos Scale | 6.94±1.67 |

Continuous variables are presented as mean±standard deviation (SD). Categorical variables are expressed as frequency and percentage (n, %). MMSE: Mini-Mental State Examination; MoCA: Montreal Cognitive Assessment; GCS: Glasgow Coma Scale.

## Complications

Spasticity was the most prevalent complication (64%), followed by joint contractures and speech disorders (42% each), dysphagia (34%), post-traumatic epilepsy (24%), neurogenic bladder (18%), and neurogenic bowel dysfunction (12%). Contracture was significantly more frequent in the severe TBI group (p=0.046).

## Functional and cognitive outcomes

Cognitive assessments showed no statistically significant differences between severity groups for MMSE, MoCA, or Rancho Los Amigos scores (all p>0.05; Table 2). Post-hoc power analysis indicated that the study had limited power to detect between-group differences of the magnitude observed (achieved power: 11.8% for MMSE, 24.1% for MoCA, 43.9% for Rancho Los Amigos, at $\alpha$=0.05), so these non-significant findings should be interpreted with caution given the substantial risk of Type II error rather than as evidence of true equivalence between groups. Descriptively, FAS scores shifted post-rehabilitation: FAS 0 decreased from 58 to 42%, while combined FAS 3–5 increased from 24 to 46% (Table 3); the p-values reported in Table 3 reflect between-group (severity) comparisons at each timepoint rather than within-group pre-post comparisons. Within-group pre-post change was formally tested with a Wilcoxon signed-rank test, which confirmed a statistically significant improvement in FAS from admission to discharge (21 patients improved, 0 worsened, 29 unchanged;

**Table 2.** Comparison of clinical variables between brain injury severity groups.

| Variable | Severe TBI (n=36, 72%) | Moderate TBI (n=8, 16%) | Mild TBI (n=6, 12%) |
|---|---|---|---|
| **Post-traumatic amnesia, n (%)** | | | **p-value** |
| None | 6 (16.7) | 6 (75.0) | 6 (100) <0.001 |
| <1 h | 0 | 0 | 0 |
| 1–24 h | 0 | 0 | 0 |
| >24 h | 30 (83.3) | 2 (25.0) | 0 |
| MMSE | 22.9±4.9 | 24.6±3.7 | 23.3±3.8 0.627 |
| Rancho Los Amigos Scale | 6.7±1.5 | 7.8±2.3 | 7.7±0.5 0.123 |
| MoCA | 18.5±6.1 | 21.6±3.5 | 20.3±3.4 0.312 |

Values are presented as mean±SD for continuous variables and as frequency and percentage (n, %) for categorical variables. Categorical variables were compared using chi-square test; continuous variables using one-way ANOVA. MMSE: Mini-Mental State Examination; MoCA: Montreal Cognitive Assessment; TBI: traumatic brain injury. p<0.05 was considered statistically significant.

**Table 3.** Comparison of Functional Ambulation Scale scores before and after treatment across brain injury severity groups.

| FAS score | Group 1 GCS 3–8 (n=36) | Group 2 GCS 9–12 (n=8) | Group 3 GCS 13–15 (n=6) |
|---|---|---|---|
| **Before treatment** | | | **p-value** |
| 0 | 21 (58.2) | 3 (37.5) | 5 (83.4) 0.060 |
| 1 | 5 (13.8) | 0 (0) | 1 (16.6) |
| 2 | 1 (2.0) | 2 (25.0) | 0 (0) |
| 3 | 9 (25.0) | 2 (25.0) | 0 (0) |
| 4 | 0 (0) | 1 (12.5) | 0 (0) |
| **After treatment** | | | **p-value** |
| 0 | 17 (47.4) | 2 (25.0) | 2 (33.0) 0.277 |
| 1 | 2 (5.5) | 0 (0) | 0 (0) |
| 2 | 1 (2.7) | 1 (12.5) | 2 (33.0) |
| 3 | 11 (30.5) | 4 (50.0) | 1 (17.0) |
| 4 | 4 (11.0) | 0 (0) | 0 (0) |
| 5 | 1 (2.7) | 1 (12.5) | 1 (17.0) |

Values are presented as frequency and percentage (n, %). FAS scores range from 0 (non-ambulatory) to 5 (normal ambulation). The p-values shown reflect chi-square comparisons across the three severity groups at each timepoint (before and after treatment, respectively); within-group pre-post change (admission vs. discharge, whole cohort) was tested separately using the Wilcoxon signed-rank test (Z=-4.078, p<0.001; 21 patients improved, 0 worsened, 29 unchanged). GCS: Glasgow Coma Scale; FAS: Functional Ambulation Scale.

Z=-4.078, p<0.001). In the severe group, FAS 0 decreased from 58.2 to 47.4%, and FAS 3–5 rose from 25 to 44.2%. In the moderate group, 50% reached FAS 3 post-treatment; in the mild group, FAS 0 decreased from 83.4 to 33%. To assess whether pooling traumatic and non-traumatic etiologies affected the cognitive and muscle strength outcomes, an independent-samples t-test compared traumatic (n=38) and non-traumatic (n=12) patients on upper extremity strength, lower extremity strength, MMSE, MoCA, and Rancho Los Amigos scores. No significant differences were found between etiology groups on any measure [upper extremity strength: t(48)=0.410, p=0.683; lower extremity strength: t(48)=-0.294, p=0.770; MMSE: t(48)=0.456, p=0.650; MoCA: t(48)=0.097, p=0.923; Rancho Los Amigos: t(48)=-0.734, p=0.467], with small effect sizes throughout (Cohen's d range 0.03–0.24), suggesting that analyzing traumatic and non-traumatic ABI together did not materially affect these outcome comparisons. Chi-square tests comparing traumatic and non-traumatic patients on complication frequency and functional status likewise showed no significant differences: spasticity [$\chi^2(1)=0.049$, p=0.825], contracture [$\chi^2(1)=0.001$, p=0.979], neurogenic bladder [$\chi^2(1)=0.524$, p=0.469], neurogenic bowel [$\chi^2(1)=2.527$, p=0.112], admission FAS [$\chi^2(4)=1.958$, p=0.743], and discharge FAS [$\chi^2(5)=4.574$, p=0.470].

## DISCUSSION

Traumatic brain injury (TBI) continues to be a significant public health concern due to its high potential for long-term disability and the complex range of physical, cognitive, and psychosocial impairments it may cause. This study offers a comprehensive evaluation of clinical, functional, and cognitive characteristics in patients with acquired brain injury, classified according to initial injury severity. One of the key strengths of the present research lies in its multidimensional approach, incorporating not only demographic and injury-related variables but also detailed assessments of complications, cognitive functioning, and post-treatment functional outcomes. By stratifying patients into severity-based subgroups and conducting within-group and between-group analyses, the study provides valuable insights into how injury severity correlates with outcomes such as post-traumatic amnesia, spasticity, contracture, and frontal lobe functioning. The longitudinal comparison of FAS scores before and after treatment underscores the dynamic nature of recovery, revealing a notable shift toward higher FAS scores in all groups, highlighting the importance of structured follow-up and tailored rehabilitation strategies even for patients with mild to moderate presentations. The findings indicate that the majority of patients with severe TBI experienced higher rates of complications such as post-traumatic epilepsy, spasticity, and joint contractures, consistent with prior literature. Post-traumatic amnesia lasting more than 24 h was most prevalent in the severe TBI group, whereas no cases of PTA were observed in mild TBI. Dysphagia was also notable in moderate and severe cases. Interestingly, deep vein thrombosis (DVT) was reported only in the mild group, which may reflect coincidental findings or limitations in documentation. In the comparison of cognitive assessments among patients with varying severity of brain injury, no statistically significant differences were found in terms of MMSE, Rancho Los Amigos Scale, or MoCA scores. Although the severe group exhibited lower mean scores in all three evaluations, these differences did not reach statistical significance. This suggests a trend toward more impaired cognitive and behavioral functioning in patients with more severe injury, but the small sample size and variability within groups may have limited the ability to detect significant effects. Our results align with recent studies emphasizing the importance of multidisciplinary rehabilitation approaches in TBI. For example, Varuges et al. demonstrated statistically significant improvements in walking distance after inpatient rehabilitation in a similar population[14]. In this study, a high prevalence of neuromuscular and neurological complications was observed. Spasticity emerged as the most common complication, affecting 64% of patients, aligning with previous reports suggesting that up to two-thirds of individuals with moderate to severe brain injury experience spasticity during recovery[15]. Joint contractures and speech disorders were also notably frequent, underscoring the need for early and sustained physiotherapeutic and speech-language interventions[16]. Dysphagia, seen in 34% of cases, is a well-documented consequence of TBI associated with increased risk of aspiration pneumonia and prolonged hospital stays[17]. Post-traumatic epilepsy affected 24% of patients, consistent with prior findings reporting an incidence of 10–25% in moderate to severe TBI[18]. Neurogenic bladder and bowel dysfunction were present in 18 and 12% of patients, respectively, reflecting common autonomic disturbances that accompany more extensive brain injuries[19]. Rare but clinically significant complications such as hydrocephalus and DVT were each identified in 2%. These findings emphasize the necessity of comprehensive, multidisciplinary rehabilitation programs capable of addressing a wide spectrum of secondary complications. It is also important to consider that spasticity and joint contractures may limit functional gains despite cognitive recovery, underscoring the need for early spasticity management and joint protection strategies. Previous studies have consistently demonstrated that PTA duration is a strong indicator of TBI severity and a reliable predictor of long-term cognitive and functional recovery. Ponsford

et al.[20] reported that longer PTA is significantly associated with poorer neuropsychological outcomes. Similarly, Walker et al.[21] demonstrated that PTA duration was the most important injury-related predictor of long-term global functional outcome[22]. In the present study, 83.3% of severe TBI patients experienced PTA longer than 24 h, compared to 25% in the moderate and 0% in the mild groups, reinforcing the utility of PTA as a marker of injury severity. Despite this significant difference, our findings did not show statistically significant between-group differences in cognitive and functional outcomes. These results may be attributed to early timing of assessments, limited subgroup sizes, and the homogenizing effect of a standardized, intensive inpatient rehabilitation program. Prior research has indicated that cognitive improvements during the inpatient phase may be similar across severity groups when structured multidisciplinary rehabilitation is provided[23]. Andelic et al.[12] reported a high prevalence of post-traumatic complications even 10 years post-injury, including spasticity, joint contractures, and neurogenic dysfunctions. Similarly, Ponsford et al.[20] found that long-term outcomes are often influenced by initial injury severity, but variability in individual recovery trajectories remains significant. The presence of dysphagia and speech disorders across all TBI groups underscores the importance of comprehensive interdisciplinary management regardless of injury severity. The significantly higher incidence of joint contractures in the severe TBI group highlights the need for proactive contracture prevention strategies during inpatient rehabilitation. The present study evaluated the impact of inpatient rehabilitation on ambulation in patients stratified by GCS. While all groups exhibited improvements in FAS scores post-treatment, these did not reach statistical significance. Al-Jarrah et al.[24] found that patients with severe and moderate TBI showed significant improvements in FIM scores following rehabilitation, whereas those with mild TBI did not, suggesting that injury severity may influence the degree of functional recovery achievable. Bilgin et al.[25] reported that GCS scores correlated with initial functional status but were not strong predictors of discharge outcomes, underscoring the importance of comprehensive assessments and individualized rehabilitation plans. Future studies with larger cohorts and longer follow-up periods are necessary to elucidate the full impact of rehabilitation on ambulation and other functional domains. This study offers several strengths. The categorization of patients based on GCS scores into severity subgroups enabled detailed analysis of injury severity–outcome relationships. The inclusion of multiple standardized scales—FAS, MMSE, MoCA, and Rancho Los Amigos—allowed for multidimensional assessment of cognitive and motor functions. Conducting the study in a specialized neurorehabilitation center ensured consistency in treatment protocols, strengthening internal validity. However, its retrospective and single-center design limits generalizability and introduces potential bias. Of the 350 patients initially screened, only 50 (14.3%) met inclusion criteria (128 excluded for missing admission/discharge assessment data, 130 for other incomplete chart data, and 42 for premature discharge due to an acute medical emergency); systematic differences between included and excluded patients—particularly the possibility that those discharged early for medical instability differed in injury severity or complication burden from those completing the program—cannot be ruled out and may limit the representativeness of the study sample. A formal comparison of included versus excluded patients on baseline demographic or clinical variables was not feasible, since the majority of exclusions were themselves due to missing or incomplete chart data for these same variables. The small subgroup sizes substantially reduced statistical power, as confirmed by post-hoc power analysis (achieved power ranging from 11.8 to 43.9% across cognitive outcomes), such that the non-significant between-group findings should be regarded as inconclusive rather than as evidence of no difference. Traumatic and non-traumatic etiologies were analyzed together given the limited sample size (n=38 vs. n=12); post-hoc comparisons found no significant differences between these groups on muscle strength, MMSE, MoCA, or Rancho Los Amigos scores, nor on complication frequency (spasticity, contracture, neurogenic bladder or bowel) or admission/discharge FAS distribution (all $p>0.05$, small effect sizes), suggesting this pooling did not materially affect the reported outcomes, though the small non-traumatic subgroup limits the power to detect subtler etiology-specific patterns. The pre-post change in FAS scores was formally tested with a Wilcoxon signed-rank test ($Z=-4.078$, $p<0.001$), confirming a statistically significant improvement; however, this whole-cohort result was not stratified by severity subgroup given the small moderate (n=8) and mild (n=6) group sizes, which limits inference about whether the improvement was uniform across severity levels. Additionally, assessments were limited to the inpatient period, precluding evaluation of long-term recovery trajectories, and the absence of a non-rehabilitation or alternative-treatment comparison group means that observed functional improvements cannot be firmly attributed to the intervention rather than to spontaneous recovery. Although session frequency and duration (5 days/week, 45 min/session) were standardized across all patients, length of stay varied considerably (14–98 days), so total treatment dose was not constant and was not analyzed as a potential confounder. Future prospective, multicenter, controlled studies with larger populations, a priori power calculations, and long-term follow-up are needed.

## CONCLUSION

This study provides valuable insights into the clinical characteristics, complications, and functional outcomes of ABI patients across varying injury severity levels. The findings underscore the importance of early, multidisciplinary rehabilitation and individualized treatment plans. Although statistical significance was not observed in all outcomes, meaningful shifts in functional status were observed post-rehabilitation. Future prospective studies with larger samples and longer follow-up are needed to elucidate long-term recovery patterns in this population.

## ETHICAL APPROVAL

All procedures were conducted in accordance with the 1964 Declaration of Helsinki. Approved by the Ethics Committee of Gaziler Physical Medicine and Rehabilitation Training and Research Hospital (Date: 26.10.2023; Protocol No: 29/008).

## INFORMED CONSENT

Informed consent was obtained from all participants.

## AUTHORS' CONTRIBUTIONS

**SGA:** Conceptualization, Data curation, Writing – original draft. ŞBS: Formal Analysis, Methodology. **SUK:** Data curation,Writing – review & editing. **KK:** Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Guan B, Anderson DB, Chen L, Feng S, Zhou H. Global, regional and national burden of traumatic brain injury and spinal cord injury, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. BMJ Open. 2023;13(10):e075049. https://doi.org/10.1136/bmjopen-2023-075049
2. Barbiellini Amidei C, Salmaso L, Fedeli U, Saia M. Traumatic brain injury incidence and mortality: a large population-based study. J Neurol. 2024;271(7):4430-40. https://doi.org/10.1007/s00415-024-12386-1
3. Selassie AW, Zaloshnja E, Langlois JA, Miller T, Jones P, Steiner C. Incidence of long-term disability following traumatic brain injury hospitalization, United States, 2003. J Head Trauma Rehabil. 2008;23(2):123-31. https://doi.org/10.1097/01.HTR.0000314531.30401.39
4. Dewan MC, Rattani A, Gupta S, Baticulon RE, Hung YC, Punchak M, et al. Estimating the global incidence of traumatic brain injury. J Neurosurg. 2018;130(4):1080-97. https://doi.org/10.3171/2017.10.JNS17352
5. Chequer Souza J, Dobson GP, Lee CJ, Letson HL. Epidemiology and outcomes of brain trauma in rural and urban populations: a systematic review and meta-analysis. Brain Inj. 2024;38(12):953-76. https://doi.org/10.1080/02699052.2024.2361641
6. Giustini A, Pistarini C, Pisoni C. Traumatic and nontraumatic brain injury. Handb Clin Neurol. 2013;110:401-9. https://doi.org/10.1016/B978-0-444-52901-5.00034-4
7. Karaboue MAA, Ministeri F, Sessa F, Nannola C, Chisari MG, Cocimano G, et al. Traumatic brain injury as a public health issue: epidemiology, prognostic factors and useful data from forensic practice. Healthcare (Basel). 2024;12(22):2266. https://doi.org/10.3390/healthcare12222266
8. Maas AIR, Menon DK, Manley GT, Abrams M, Åkerlund C, Andelic N, et al. Traumatic brain injury: progress and challenges in prevention, clinical care, and research. Lancet Neurol. 2022;21(11):1004-60. https://doi.org/10.1016/S1474-4422(22)00309-X
9. Bradley L, Wheelwright S. The impact of delays in transfer to specialist rehabilitation on outcomes in patients with acquired brain injury. Clin Rehabil. 2024;38(11):1552-8. https://doi.org/10.1177/02692155241284866
10. Khan F, Baguley IJ, Cameron ID. 4: Rehabilitation after traumatic brain injury. Med J Aust. 2003;178(6):290-5. https://doi.org/10.5694/j.1326-5377.2003.tb05199.x
11. Turner-Stokes L, Pick A, Nair A, Disler PB, Wade DT. Multidisciplinary rehabilitation for acquired brain injury in adults of working age. Cochrane Database Syst Rev. 2015;2015(12):CD004170. https://doi.org/10.1002/14651858.CD004170.pub3
12. Andelic N, Howe EI, Hellstrøm T, Sanchez MF, Lu J, Løvstad M, et al. Disability and quality of life 20 years after traumatic brain injury. Brain Behav. 2018;8(7):e01018. https://doi.org/10.1002/brb3.1018
13. Shen Y, Jiang L, Lai J, Hu J, Liang F, Zhang X, et al. A comprehensive review of rehabilitation approaches for traumatic brain injury: efficacy and outcomes. Front Neurol. 2025;16:1608645. https://doi.org/10.3389/fneur.2025.1608645
14. Varuges JA, Mazlan M. Rehabilitation characteristics and outcomes of adults with traumatic brain injury: a retrospective study in UMMC, a tertiary centre in Klang Valley. Med J Malaysia. 2023;78(2):190-6. PMID: 36988529
15. Zandalasini M, Pelizzari L, Ciardi G, Giraudo D, Guasconi M, Paravati S, et al. Bowel dysfunctions after acquired brain injury: a scoping review. Front Hum Neurosci. 2023;17:1146054. https://doi.org/10.3389/fnhum.2023.1146054
16. Raciti L, Raciti G, Pulejo G, Conti-Nibali V, Calabrò RS. Neurogenic dysphagia and nutrition in disorder of consciousness: an overview with practical advices on an "Old" but still actual clinical problem. Medicines (Basel). 2022;9(2):16. https://doi.org/10.3390/medicines9020016
17. Eskildsen SJ, Jakobsen D, Riberholt CG, Poulsen I, Curtis DJ. Protocol for a scoping review study to identify and map treatments for dysphagia following moderate to severe acquired brain injury. BMJ Open. 2019;9(7):e029061. https://doi.org/10.1136/bmjopen-2019-029061

18. Pease M, Gonzalez-Martinez J, Puccio A, Nwachuku E, Castellano JF, Okonkwo DO, et al. Risk factors and incidence of epilepsy after severe traumatic brain injury. Ann Neurol. 2022;92(4):663-9. https://doi.org/10.1002/ana.26443

19. Pelizzari L, Antoniono E, Giraudo D, Ciardi G, Lamberti G. Fecal incontinence after severe brain injury: a barrier to discharge after inpatient rehabilitation? Neurol Int. 2023;15(4):1339-51. https://doi.org/10.3390/neurolint15040084

20. Ponsford J, Draper K, Schönberger M. Functional outcome 10 years after traumatic brain injury: its relationship with demographic, injury severity, and cognitive and emotional status. J Int Neuropsychol Soc. 2008;14(2):233-42. https://doi.org/10.1017/S1355617708080272

21. Walker WC, Stromberg KA, Marwitz JH, Sima AP, Agyemang AA, Graham KM, et al. Predicting long-term global outcome after traumatic brain injury: development of a practical prognostic tool using the traumatic brain injury model systems national database. J Neurotrauma. 2018;35(14):1587-95. https://doi.org/10.1089/neu.2017.5359

22. Vile AR, Jang K, Gourlay D, Marshman LAG. Posttraumatic amnesia: a systematic review and meta-analysis. proposal for a new severity classification. World Neurosurg. 2022;162:e369-93. https://doi.org/10.1016/j.wneu.2022.03.018

23. Pommerich UM, Stubbs PW, Eggertsen PP, Fabricius J, Nielsen JF. Regression-based prognostic models for functional independence after postacute brain injury rehabilitation are not transportable: a systematic review. J Clin Epidemiol. 2023;156:53-65. https://doi.org/10.1016/j.jclinepi.2023.02.009

24. Al-Jarrah MD, Nazzal ME, Jamous MA, Azab MA, Maayah MF. Association between the functional independence measure and Glasgow Coma Scale regarding the rehabilitation outcomes of traumatic brain injury. Neurosciences (Riyadh). 2009;14(1):41-4. PMID: 21048572

25. Bilgin S, Guclu-Gunduz A, Oruckaptan H, Kose N, Celik B. Gait and Glasgow Coma Scale scores can predict functional recovery in patients with traumatic brain injury. Neural Regen Res. 2012;7(25):1978-84. https://doi.org/10.3969/j.issn.1673-5374.2012.25.009

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260577

# Cancer-related deaths among cases referred for forensic evaluation in Türkiye: a 5-year retrospective analysis

Ahmet Burak Aydemir[1], Veysel Osman Soğukpınar[1]*, Humman Şen[2], Ferah Karayel[2]

## SUMMARY

**OBJECTIVE:** Cancer is a major global cause of mortality and holds particular importance in forensic cases regarding diagnosis, treatment processes, and determination of cause of death.

**METHODS:** This retrospective study analyzed 163 cancer-related death cases referred for forensic evaluation to the First Specialized Board of the Council of Forensic Medicine of Türkiye between 2019 and 2023. Demographic, clinical, pathological, and forensic data were evaluated comprehensively.

**RESULTS:** Most cases were male (79.1%), with a mean age of 57.3±18.5 years. The most common primary cancer was lung (30.7%). Diagnosis relied mainly on medical records (84%), while 12.3% were diagnosed solely by autopsy. Comorbidities were frequent (74.2%), the metastasis rate was high (46.0%), and deaths mostly occurred in hospitals (78.5%). Malpractice claims were present in 36.8%, particularly higher in interventional treatments.

**CONCLUSION:** Autopsy and histopathological examinations remain essential for diagnostic accuracy in forensic cancer-related deaths. Malpractice allegations were more frequent among cases undergoing surgery and chemotherapy/radiotherapy, although these findings reflect allegations rather than confirmed malpractice. A multidisciplinary approach integrating clinical, pathological, and forensic data is crucial for medicolegal evaluation.

**KEYWORDS:** Autopsy. Neoplasms. Cause of death. Malpractice. Neoplasm metastasis.

## INTRODUCTION

Cancer remains one of the leading causes of morbidity and mortality worldwide. Current data indicate an increasing cancer burden both globally and in Türkiye. According to GLOBOCAN (Global Cancer Observatory: Cancer Today) 2022 data, millions of new cancer cases and related deaths are reported worldwide, and cancer ranks among the leading causes of death in Türkiye[1-3].

The forensic medical evaluation of cancer cases is particularly important in cases involving sudden and unexpected deaths, allegations of negligence or malpractice related to diagnostic and therapeutic processes, and situations requiring clarification of the cause of death. Findings on malignancy detection in forensic autopsy series vary in the literature, mainly due to differences in study design, autopsy indications, and histopathological examination approaches[4,5]. Therefore, standardized autopsy protocols and detailed histopathological examinations are critical for accurate judicial processes and reliable public health data[6].

In forensic practice, cancer cases are classified into three groups: primary malignancies presenting as sudden deaths diagnosed at autopsy; suspicious deaths in individuals with known cancer; and cases involving malpractice allegations. Each requires specific evaluation and a multidisciplinary approach[7].

This study aimed to retrospectively analyze demographic, clinical, and pathological findings, as well as forensic evaluation processes, in cancer-related deaths over five years in Türkiye, and to contribute to forensic practice.

## METHODS

### Case selection and data collection

This retrospective study included 163 forensic cases evaluated by the First Specialized Board of the Council of Forensic Medicine of the Republic of Türkiye between January 1, 2019, and December 31, 2023, in which the cause of death was determined as cancer. Because cases were included according to the

[1]Ministry of Justice, Council of Forensic Medicine, The Group Presidency of Forensic Medicine – Van, Türkiye.
[2]Ministry of Justice, The Council of Forensic Medicine – Istanbul, Türkiye.
*Corresponding author: v.o.s23@hotmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on May 17, 2026. Accepted on July 11, 2026.
Scientific Editor: José Maria Soares Júnior

date of forensic board evaluation rather than the date of death, some deaths had occurred in earlier years. Cases were selected according to definitive causes of death established by the board after evaluation of forensic files, while those referred for insurance assessment were excluded. Cases referred solely for insurance assessment were excluded because they follow a distinct legal evaluation pathway and were not considered comparable to medicolegal death investigations.

Cases were analyzed in terms of demographic characteristics (sex, age), place and year of death, geographic distribution, hospitalization history, and the interval between the last hospital admission and death. Primary cancer type, diagnostic methods (medical records, statements, autopsy findings), treatment modalities, comorbid systemic diseases, and autopsy findings were also evaluated. Histopathological results, including tumor presence, involved organs and systems, and pathological diagnoses, were recorded. Additionally, metastasis status and localization, toxicological findings, definitive causes of death, and malpractice allegations were assessed. Toxicological categories were based on all substances detected in biological samples. Consequently, chronic disease medications and cardiopulmonary resuscitation drugs could be recorded concurrently within the same case. All variables were analyzed holistically by integrating clinical, pathological, and forensic data. For cause-of-death classification, cases categorized as "cancer and infection/multiple organ failure" were those in which infection or multiple organ failure was identified as the immediate terminal event, whereas "primary cancer and complications" referred to progressive cancer-related mortality without a dominant infectious process.

Histopathological data included both postmortem tissue examinations and available antemortem pathology reports (biopsy and surgical pathology specimens) contained within the forensic case files. Therefore, the number of cases with histopathological information may slightly exceed the number of autopsied cases.

### Data analysis

Data were analyzed using IBM SPSS Statistics v.27. Descriptive statistics were presented as number (n) and percentage (%) for categorical variables, and mean±standard deviation and median (minimum–maximum) for continuous variables. Chi-square and Fisher's exact tests were used for categorical comparisons; distribution was assessed by the Kolmogorov-Smirnov test. Student's t-test or Mann-Whitney U test was applied, and Kruskal-Wallis was used for multiple groups. Statistical significance was set at $p<0.05$. Because the study was exploratory in nature, reported p-values were not adjusted for multiple comparisons and should be interpreted cautiously.

### Ethics approval

Approval for this study was obtained from the Education and Scientific Research Commission of the Council of Forensic Medicine (decision dated April 30, 2024; No. 21589509/2024/323), and the principles of the Declaration of Helsinki were observed.

## RESULTS

Of the 163 cases, 79.1% (n=129) were male. Cases were concentrated in older age groups, most frequently in the 60–69 range (28.2%; n=46). The mean age was 57.26±18.5 years. Most deaths occurred in hospitals (78.5%; n=128), with increased cases between 2020 and 2022. Autopsy was performed in 81.6% (n=133), and 71.8% (n=117) had prior hospitalization. The interval between last admission and death was most often 8–30 days (25.8%; n=42) (Table 1).

The most common primary cancer was lung (30.7%; n=50), followed by small and large intestine (9.2%; n=15), liver (7.4%; n=12), and hematological malignancies (6.7%; n=11); breast cancer accounted for 3.1% (n=5). Diagnosis relied mainly on medical records (84%; n=137), followed by statements (70.6%; n=115) and autopsy findings (48.5%; n=79). In 12.3% (n=20), diagnosis was established solely by autopsy.

Medical therapy alone was most common (26.4%; n=43), with combined treatments also frequent (Table 2). At least one systemic disease was present in 74.2%. Cardiovascular diseases (52.2%; n=85) were most frequent, followed by diabetes mellitus (18.4%; n=30) and chronic obstructive pulmonary disease (COPD) (17.2%; n=28). No or unknown comorbidity was present in 25.8% (n=42). At autopsy, mass lesions and serous fluid accumulation (39.3%; n=64) predominated; no macroscopic findings were observed in 12.3% (n=20), and prior organ removal in 4.9% (n=8).

Histopathological examination was performed in 82.8% (n=135), with tumors detected in 46.6% (n=76). The lung was most frequently involved (17.8%; n=29), followed by liver (6.1%; n=10), bone marrow (3.7%; n=6), and pancreas (3.1%; n=5). The respiratory system predominated (19%; n=31) (Table 2). The most common diagnosis was malignant tumor (24.5%; n=40), followed by adenocarcinoma (7.4%; n=12) and squamous cell carcinoma (3.7%; n=6).

Metastasis was identified in 46.0% (n=75), most commonly in the liver (22.1%; n=36), followed by brain (10.4%; n=17) and heart (9.2%; n=15). Toxicology most frequently showed

**Table 1.** Demographic and clinical characteristics of the cases.

| Characteristics | | n | % |
|---|---|---|---|
| Sex | Male | 129 | 79.1 |
| | Female | 34 | 20.9 |
| Age groups (decades) | 0–9 | 7 | 4.3 |
| | 10–19 | 0 | 0.0 |
| | 20–29 | 7 | 4.3 |
| | 30–39 | 8 | 4.9 |
| | 40–49 | 16 | 9.8 |
| | 50–59 | 39 | 23.9 |
| | 60–69 | 46 | 28.2 |
| | 70–79 | 26 | 16.0 |
| | 80–89 | 12 | 7.4 |
| | 90–99 | 2 | 1.2 |
| Place of death | Hospital | 128 | 78.5 |
| | Home | 33 | 20.3 |
| | Nursing home/care facility | 2 | 1.2 |
| Year of death | 2016 | 2 | 1.2 |
| | 2017 | 6 | 3.7 |
| | 2018 | 20 | 12.3 |
| | 2019 | 22 | 13.5 |
| | 2020 | 28 | 17.2 |
| | 2021 | 31 | 19.0 |
| | 2022 | 37 | 22.7 |
| | 2023 | 17 | 10.4 |
| Geographical regions | Marmara | 59 | 36.2 |
| | Aegean | 17 | 10.4 |
| | Mediterranean | 14 | 8.6 |
| | Central anatolia | 41 | 25.2 |
| | Black sea | 12 | 7.4 |
| | Eastern anatolia | 6 | 3.7 |
| | Southeastern anatolia | 14 | 8.6 |
| Autopsy status | Performed | 133 | 81.6 |
| | Not performed | 28 | 17.2 |
| | Exhumation | 2 | 1.2 |
| History of hospitalization | Present | 117 | 71.8 |
| | Absent | 46 | 28.2 |
| Time between last hospital admission and death | 0–24 h | 40 | 24.5 |
| | 2–3 days | 20 | 12.3 |
| | 4–7 days | 19 | 11.7 |
| | 8–30 days | 42 | 25.8 |
| | 1–4 months | 12 | 7.4 |
| | Non-hospital deaths | 30 | 18.4 |

n: number; %: percentage. Cases were included according to the date of forensic board evaluation (2019–2023); therefore, the year of death may precede the study evaluation period.

**Table 2.** Clinical, pathological, toxicological characteristics and outcomes of the cases.

| Characteristics | | n | % |
|---|---|---|---|
| Treatment modalities | Medical treatment only | 43 | 26.4 |
| | Unspecified/unknown | 32 | 19.6 |
| | Surgery+CT/RT+medical | 27 | 16.6 |
| | CT/RT+medical | 25 | 15.3 |
| | Surgery+medical | 13 | 8.0 |
| | Surgery only | 12 | 7.4 |
| | CT/RT only | 11 | 6.7 |
| Systems with histopathologically detected tumor | Respiratory | 31 | 19.0 |
| | Gastrointestinal | 25 | 15.3 |
| | Hematolymphoid | 5 | 3.1 |
| | Genitourinary | 7 | 4.3 |
| | Central nervous system | 2 | 1.2 |
| | Endocrine | 1 | 0.6 |
| | Other | 5 | 3.1 |
| | No tumor detected | 87 | 53.4 |
| Toxicological findings | CDD+CPR drugs | 76 | 46.6 |
| | CDD | 32 | 19.6 |
| | CPR drugs | 6 | 3.7 |
| | Substances of abuse+CPR drugs+CDD | 4 | 2.5 |
| | Chemotherapeutic agents | 2 | 1.2 |
| | No drug/substance detected | 19 | 11.7 |
| | Unknown | 24 | 14.7 |
| Definitive cause of death | Primary cancer and complications | 86 | 52.8 |
| | Cancer and infection/MOF | 45 | 27.6 |
| | Cancer and systemic diseases | 12 | 7.4 |
| | PCSC | 9 | 5.5 |
| | Other (e.g., trauma and cancer) | 8 | 4.9 |
| | UA (possible cancer) | 2 | 1.2 |
| | Drug intoxication and cancer | 1 | 0.6 |
| Malpractice allegation | Present | 60 | 36.8 |
| | Absent | 103 | 63.2 |

CT: chemotherapy; RT: radiotherapy; CDD: chronic disease drugs; CPR: cardiopulmonary resuscitation; MOF: multiple organ failure; PCSC: primary cancer surgery and complications; UA: undetermined at autopsy; n: number; %: percentage.

combined chronic disease and resuscitation drugs (46.6%; n=76). The most common cause of death was primary cancer and its complications (52.8%; n=86), and malpractice allegations were present in 36.8% (n=60) (Table 2).

Lung cancer was more frequent in males (32.6%; n=42) than females (23.5%; n=8), with a significant association between cancer type and sex (p=0.016), but not age (p=0.196). Respiratory system cancers were more frequent in those ≥60 years (33.1%; n=54), without significance (p=0.106). No associations were found between systems and metastasis (p=0.260) or causes of death (p=0.303). Medical therapy was more frequent in cases without histopathologically detected tumors (75.9%; n=66). Surgical treatment was significantly associated with histopathological tumor detection (p=0.014), whereas no significant associations were observed for chemotherapy/radiotherapy (p=0.093) or medical treatment (p=0.139). Malpractice allegations were more frequent among cases undergoing surgery (48.1%; n=25) and chemotherapy/radiotherapy (47.6%; n=30) than among those who did not receive these treatments (p=0.041 and p=0.023, respectively), whereas no significant association was observed for medical treatment alone (p=0.265).

## DISCUSSION

This study demonstrates a clear relationship between clinical processes and forensic medical evaluation in cancer-related deaths, indicating a multidimensional interaction between clinical, pathological, and forensic components.

Most cases were male (79.1%) and in older age groups, with a mean age of 57.26±18.5 years. Similarly, Çengel et al.[8] reported a male proportion of 79.7% in malignancy cases. This male predominance is consistent with cancer epidemiology, as GLOBOCAN 2022 data show higher incidence rates in men both globally and in Türkiye[9]. These findings highlight the importance of demographic characteristics in forensic evaluation.

The predominance of hospital deaths and limited nursing home deaths suggests insufficient institutional care for terminal cancer patients. This aligns with reports indicating limited institutional care capacity in Türkiye[10] and high care needs in nursing home populations[11]. The increase in cases during 2020–2022 may reflect disruptions in cancer care during the COVID-19 pandemic, including delays in diagnosis and treatment[12], potentially leading to advanced-stage disease and increased forensic case burden. These findings underscore the need to strengthen palliative and end-of-life care services.

The high rate of hospitalization and the concentration of deaths within early and intermediate periods after admission suggest the importance of clinical complications. The coexistence of cancer and infection, along with the predominance of respiratory system cancers, supports the role of infections in mortality. Medical intervention-related complications, particularly lower respiratory tract infections, are more frequent in hospital deaths[13]. Thus, mortality in cancer patients is influenced not only by disease progression but also by complications arising during clinical care, emphasizing the need for comprehensive evaluation of healthcare history.

Within this forensic referral cohort, lung cancer was the most frequently identified primary malignancy, and respiratory and gastrointestinal system tumors predominated. The higher frequency of lung cancer among men was broadly consistent with epidemiological observations reported in the literature[9]. However, because the study population consisted exclusively of cases referred for medicolegal evaluation, these findings should not be interpreted as representative of the overall distribution of cancer-related deaths in Türkiye. The relatively low frequency of breast cancer among women may reflect the forensic characteristics of the study population. Histopathological findings, with the lung as the most frequently involved organ and the respiratory system predominating, support this distribution. The absence of tumor findings in more than half of cases may be related to widespread metastasis and organ failure in terminal stages. Diagnosis based mainly on medical records, with a notable proportion established only at autopsy, highlights the limitations of clinical evaluation and the importance of autopsy and histopathology[14,15]. These findings emphasize the necessity of integrated evaluation for diagnostic accuracy.

Multimodal treatment approaches are emphasized in oncology, with therapy guided by tumor characteristics[16]. In this study, medical therapy was most common, while combined treatments were also frequent. The significant association between surgical treatment and histopathological diagnosis indicates concordance between treatment and diagnostic processes. However, missing treatment data in 19.6% of cases underscores the importance of complete medical documentation, as incomplete records may negatively affect forensic processes[17]. Systematic documentation is therefore critical in both clinical and forensic contexts.

Comorbid diseases are common in cancer and significantly influence prognosis[18,19]. In this study, at least one systemic disease was present in approximately three-quarters of cases, consistent with the literature. The higher rate compared to the 48% reported by Karakoç[20] may be due to inclusion of only fatal cases. Cardiovascular diseases and diabetes mellitus were the most common comorbidities, consistent with Renzi et al.[18], while respiratory diseases such as COPD support their negative impact, particularly in lung cancer cases[19].

Macroscopic examination is a fundamental step in autopsy diagnosis, although findings may be limited, requiring

histopathological confirmation[21]. The detection of mass lesions and serous fluid accumulation, alongside cases without macroscopic findings, supports this. Prior surgical removal of organs may limit macroscopic evaluation, necessitating correlation with clinical history. The predominance of malignant tumors, adenocarcinoma, and squamous cell carcinoma aligns with the literature, particularly for lung cancer subtypes[22]. The most frequent metastasis sites—liver, brain, and heart—are consistent with organotropic spread patterns[23]. The relatively high frequency of cardiac metastases should be interpreted cautiously. Cardiac metastases are often underdiagnosed during life and may be detected more frequently during postmortem examinations, particularly in advanced-stage disease. However, the lack of significant association between systems and metastasis suggests a multifactorial process. These findings reinforce the importance of integrating macroscopic, histopathological, and clinical data.

In this study, toxicological analysis most frequently revealed the combined presence of medications for chronic diseases and cardiopulmonary resuscitation drugs. Additionally, the low detection rate of chemotherapeutic agents and the absence of any detectable substances or unavailable results in some cases are noteworthy. In forensic toxicology, the analysis of biological samples is essential for identifying drugs used prior to death and evaluating their potential contribution to mortality[24]. This may be related to the postmortem stability of chemotherapeutic agents, analytical limitations, and variability associated with postmortem processes. Therefore, toxicological findings should be interpreted within the clinical and forensic context of the case.

Malpractice allegations were identified in 36.8% of cases and were significantly more frequent in those undergoing surgery and chemotherapy/radiotherapy. In oncology, malpractice claims are often associated with treatment processes and care organization[25]. However, this study evaluated malpractice allegations rather than confirmed malpractice determinations. The higher frequency of allegations among cases receiving surgery or chemotherapy/radiotherapy may reflect treatment complexity, disease severity, or patient and family expectations rather than deficiencies in care. Although complications may lead to serious outcomes, not all adverse events constitute malpractice. In fatal cases, autopsy findings remain critical for the evaluation of malpractice claims.

### Limitations

This study has several limitations. First, the study population consisted exclusively of cases referred to the First Specialized Board of the Council of Forensic Medicine, introducing referral and selection bias and limiting the generalizability of the findings. Therefore, the observed distributions of cancer types, clinical characteristics, and medicolegal findings may not reflect those of all cancer-related deaths in Türkiye. Second, treatment information was unavailable in 19.6% of cases. Third, no non-forensic comparison group was available. Fourth, the study evaluated malpractice allegations rather than substantiated or legally confirmed malpractice determinations. Fifth, toxicological findings were interpreted qualitatively because quantitative postmortem drug concentrations were not consistently available. Finally, although additional multivariable analyses may have provided further insights, the retrospective nature of the study and the lack of access to an analyzable individual-level dataset precluded further modeling beyond the predefined analyses.

## CONCLUSION

This study demonstrated that autopsy and histopathological examinations remain essential for diagnostic accuracy in forensic cancer-related deaths. Malpractice allegations were more frequent among cases undergoing surgery and chemotherapy/radiotherapy, although these findings reflect allegations rather than confirmed malpractice. The results highlight the importance of integrating clinical, pathological, and forensic data through a multidisciplinary approach in medicolegal evaluations of cancer-related deaths.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Approval for this study was obtained from the Education and Scientific Research Commission of the Council of Forensic Medicine (decision dated April 30, 2024; No. 21589509/2024/323), and the principles of the Declaration of Helsinki were observed.

## AUTHORS' CONTRIBUTIONS

**ABA:** Conceptualization, Formal Analysis, Writing – review & editing. **VOS:** Conceptualization, Formal Analysis, Writing – original draft. **HŞ:** Conceptualization, Data curation, Supervision. **FK:** Conceptualization, Methodology.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-49. https://doi.org/10.3322/caac.21660
2. Republic of Türkiye Ministry of Health, General Directorate of Health Information Systems. Health Statistics Yearbook 2022. Ankara; 2024. Publication No: 1279. [cited on 2026 Apr 1]. Available from: https://ekutuphane.saglik.gov.tr/Ekutuphane/kitaplar/siy202205042024pdf.pdf
3. Cao W, Qin K, Li F, Chen W. Comparative study of cancer profiles between 2020 and 2022 using global cancer statistics (GLOBOCAN). J Natl Cancer Cent. 2024;4(2):128-34. https://doi.org/10.1016/j.jncc.2024.05.001
4. Shojania KG, Burton EC, McDonald KM, Goldman L. Changes in rates of autopsy-detected diagnostic errors over time: a systematic review. JAMA. 2003;289(21):2849-56. https://doi.org/10.1001/jama.289.21.2849
5. Robb TJ, Tse R, Blenkiron C. Reviving the autopsy for modern cancer evolution research. Cancers (Basel). 2021;13(3):409. https://doi.org/10.3390/cancers13030409
6. Burton EC, Troxclair DA, Newman WP, 3rd. Autopsy diagnoses of malignant neoplasms: how often are clinical diagnoses incorrect? JAMA. 1998;280(14):1245-8. https://doi.org/10.1001/jama.280.14.1245
7. Saukko P, Knight B. Knight's forensic pathology. 4th ed. Boca Raton: CRC Press (Taylor & Francis Group); 2016.
8. Çengel G, Demiroğlu Uyanıker Z, Ünüvar Ü, Can İÖ, Olgun EG. The role of forensic autopsies in diagnosis of cancer. Bull Leg Med. 2013;18(2):47-51.
9. International Agency for Research on Cancer (IARC). Global cancer observatory: cancer today (GLOBOCAN 2022). Lyon (France): IARC; 2022. [cited on 2026 Apr 8]. Available from: https://gco.iarc.fr/
10. Turkish Statistical Institute (TÜİK). Türkiye Elderly Profile Survey, 2023. Ankara: TÜİK; 2024. Publication No: 53809. [cited on 2026 Apr 9]. Available from: https://veriportali.tuik.gov.tr/tr/press/53809
11. Emiral E, Değirmenci B, Aslan M, Akgüney B. Post-mortem evaluation of deaths in nursing homes and care facilities in Ankara province [Turkish]. ESTUDAM PHJ. 2024;9(3):268-75. https://doi.org/10.35232/estudamhsd.1510547
12. Cosimo S, Susca N, Apolone G, Silvestris N, Racanelli V. The worldwide impact of COVID-19 on cancer care: a meta-analysis of surveys published after the first wave of the pandemic. Front Oncol. 2022;12:961380. https://doi.org/10.3389/fonc.2022.961380
13. Haukland EC, Mevik K, Plessen C, Nieder C, Vonen B. Contribution of adverse events to death of hospitalised patients. BMJ Open Qual. 2019;8(1):e000377. https://doi.org/10.1136/bmjoq-2018-000377
14. Numaguchi A, Ishii A, Natsume J, Saitoh S, Aoki Y, Yoshikawa T, et al. Medical records as screening tools for child death review in Japan. Pediatr Int. 2023;65(1):e15692. https://doi.org/10.1111/ped.15692
15. Latten BGH, Kubat B, Brandt PA, Zur Hausen A, Schouten LJ. Cause of death and the autopsy rate in an elderly population. Virchows Arch. 2023;483(6):865-72. https://doi.org/10.1007/s00428-023-03571-0
16. Consalvo F, Simone M, Scarpa A, Acerra A, Salzano FA, Fineschi V, et al. Challenges and complications in the management of advanced oropharyngeal carcinoma: role of post-mortem diagnosis and future perspectives. J Clin Med. 2024;13(17):5198. https://doi.org/10.3390/jcm13175198
17. Cattaneo C, Tambuzzi S, Vecchi S, Maggioni L, Costantino G. Consequences of the lack of clinical forensic medicine in emergency departments. Int J Legal Med. 2024;138(1):139-50. https://doi.org/10.1007/s00414-023-02973-8
18. Renzi C, Kaushal A, Emery J, Hamilton W, Neal RD, Rachet B, et al. Comorbid chronic diseases and cancer diagnosis: disease-specific effects and underlying mechanisms. Nat Rev Clin Oncol. 2019;16(12):746-61. https://doi.org/10.1038/s41571-019-0249-6
19. Morishima T, Matsumoto Y, Koeda N, Shimada H, Maruhama T, Matsuki D, et al. Impact of comorbidities on survival in gastric, colorectal, and lung cancer patients. J Epidemiol. 2019;29(3):110-5. https://doi.org/10.2188/jea.JE20170241
20. Karakoç MD. Polypharmacy and drug interactions in cancer patients [Turkish]. J Adnan Menderes Univ Health Sci Fac. 2022;6(1):31-42. https://doi.org/10.46237/amusbfd.922779
21. Singh D, Tiwari RC, Kumar A, Bhute AR, Meshram RP, Dikshit M, et al. A comprehensive review of pathological examination in forensic medicine: past, present, and future. Cureus. 2022;14(3):e22740. https://doi.org/10.7759/cureus.22740
22. Zhang Y, Vaccarella S, Morgan E, Li M, Etxeberria J, Chokunonga E, et al. Global variations in lung cancer incidence by histological subtype in 2020: a population-based study. Lancet Oncol. 2023;24(11):1206-18. https://doi.org/10.1016/S1470-2045(23)00444-8
23. Shalata W, Abu Jama A, Abu Salman A, Golosky M, Solomon A, Abu Saleh O, et al. Unexpected and rare sites of metastasis in oncologic patients. J Clin Med. 2023;12(20):6447. https://doi.org/10.3390/jcm12206447
24. Truver MT, Chronister CW, Davis GG, Gray TR, Hartman RL, Kahl JH, et al. Application of professional best practices in postmortem forensic toxicology. J Anal Toxicol. 2025;49(8):529-41. https://doi.org/10.1093/jat/bkaf061
25. Doolin JW, Schaffer AC, Tishler RB, Jacobson JO. An analysis of medical malpractice claims against medical oncologists from a national database: implications for safer practice. J Healthc Risk Manag. 2024;43(3):18-28. https://doi.org/10.1002/jhrm.21563

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260588

# Evaluation of infants aged 0–12 months admitted to the pediatric emergency department: the example of Batman Training And Research Hospital

Semih Canpolat[1], Mehmet Kenan Nahya[2]*, Ertuğrul Dedeoğlu[2], Mesut Saka[3], Mehmet Emin Parlak[4], Aylin Aktı[5]

## SUMMARY

**OBJECTIVE:** This study aimed to assess the reasons for admission, demographic and clinical characteristics, diagnostic procedures, hospitalization rates, and repeat visits among infants aged 0–12 months transported to the pediatric emergency department by 112 Emergency Medical Services ambulances in 2023, with a focus on. The study further evaluated the urgency of visits and sought to identify potentially preventable hospital admissions, while also examining the clinical decision-making processes involved.

**METHODS:** This retrospective investigation involved infants brought to the Pediatric Emergency Department of Batman Training and Research Hospital from January 1 to December 31, 2023. Hospital records were examined for demographic reasons for admission, laboratory and imaging evaluations, hospitalization or referral outcomes, and repeat visits. Patients were grouped as neonates (0–28 days) or post-neonatal infants (29 days–12 months). Statistical analyses were conducted using Statistical Package for the Social Sciences 26.

**RESULTS:** Of 3,950 ambulance transports, 957 (24.2%) involved infants aged 0–12 months; 55.6% were male. The leading causes for admission were respiratory distress (38.8%), jaundice (21.8%), and fever (8.3%). Hospitalization was needed in 60.6% of cases, and 20.0% were admitted to neonatal intensive care. Approximately 37.2% were discharged after clinical examination and limited diagnostic workup. Repeat presentations were seen in 11.3% of patients. Most admissions occurred on weekends and during off-hours periods.

**CONCLUSION:** Many ambulance transports of infants involved non-urgent conditions, yet hospitalization rates remained high, especially among neonates. Improving parental education, scheduled follow-up for high-risk infants, and primary care access could reduce unnecessary visits.

**KEYWORDS:** Ambulance. Infant. Newborn. Emergency service, hospital.

## INTRODUCTION

The infant period is a highly vulnerable stage of life, characterized by ongoing physiological, metabolic, and environmental adaptation, with relatively higher risks of morbidity and mortality. The neonatal period, as the most delicate phase of infancy, is particularly significant, since illnesses may present with subtle clinical findings yet progress rapidly. In recent years, influenced by national health, economic, and social factors, shorter postpartum hospital stays have contributed to an increase in pediatric emergency department visits among neonates and infants, either due to emerging medical issues or for routine post-discharge evaluation[1].

Emergency conditions are those with sudden onset that may be life-threatening or carry a risk of permanent health deterioration, thus requiring prompt medical attention[2]. In infancy, however, symptoms are often mild or nonspecific, which can lead to visits for non-urgent problems. This contributes to overcrowding in pediatric emergency departments and makes efficient healthcare delivery more challenging.

In recent years, neonatal intensive care unit admission rates have been reported to be approximately 8–10%, which is thought to reflect improvements in antenatal care and more structured perinatal monitoring processes[3]. Despite these advances in healthcare delivery, mortality during the neonatal and early

[1]Ankara Kahramankazan State Hospital, Department of Pediatrics – Ankara, Turkey.
[2]Şanlıurfa Training and Research Hospital, Department of Pediatrics – Şanlıurfa, Turkey.
[3]Uşak University, Department of Pediatrics – Uşak, Turkey.
[4]University of Health Sciences, Mersin City Education and Research Hospital, Department of Child Health and Diseases – Mersin, Turkey.
[5]Istanbul University, Istanbul Faculty of Medicine, Department of Pediatrics, Division of Pediatric Nephrology – Istanbul, Turkey.
*Corresponding author: kenannahya@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on April 13, 2026. Accepted on July 11, 2026.
Scientific Editor: José Maria Soares Júnior

infant periods continues to be largely associated with acute and rapidly developing conditions that require timely recognition and immediate medical intervention. Early diagnosis, close clinical follow-up, and prompt initiation of appropriate treatment remain critical factors in improving outcomes and reducing both mortality and morbidity rates in this vulnerable population[4].

Prehospital emergency medical services, particularly 112 Emergency Medical Services ambulances, have an important role in identifying high-risk infants early and ensuring their rapid transfer to suitable healthcare facilities. Careful assessment of potential risks before discharge, along with comprehensive parental guidance and well-structured follow-up planning, can help reduce avoidable emergency department visits and repeated hospital admissions[5]. In this context, examining the reasons for admission, clinical features, and outcomes of infants aged 0–12 months transported to the pediatric emergency department by ambulance is essential for both improving neonatal health outcomes and supporting effective healthcare service planning.

## METHODS

This study was carried out within the scope of a research project conducted at Batman Training and Research Hospital and was approved by the Mardin Training and Research Hospital Scientific Research Ethics Committee (approval date: October 8, 2024; approval number: 2024/10-27). All study procedures were conducted in accordance with the principles outlined in the Declaration of Helsinki.

The study was planned as a retrospective, descriptive investigation. Data were obtained from the Hospital Information Management System of Batman Training and Research Hospital, and records of infants aged 0–12 months who were brought to the Pediatric Emergency Department by 112 Emergency Medical Services ambulances between January 1, 2023, and December 31, 2023, were reviewed in detail retrospectively.

Patients' demographic features, reasons for admission, laboratory tests and imaging studies performed, hospitalization or referral status, and repeat emergency department visits were evaluated using data from the Hospital Information Management System. Visits by the same patient at different times were recorded as separate entries, and repeat admissions were assessed as a distinct variable during analysis.

During the study period, a total of 3,950 patients transported to the pediatric emergency department by 112 Emergency Medical Services ambulances were identified through the Hospital Information Management System. Among them, 957 infants aged 0–12 months formed the study group. No sampling method was applied, and all eligible cases were included directly in the analysis.

### Statistical analysis

Statistical analyses were performed using Statistical Package for the Social Sciences version 26. Continuous variables were expressed as mean±standard deviation or median (minimum–maximum), depending on their distribution. Categorical variables were presented as numbers (n) and percentages (%). Descriptive statistics were used to evaluate the demographic and clinical characteristics of the study population.

## RESULTS

In 2023, a total of 3,950 pediatric patients were transported to the pediatric emergency department by 112 Emergency Medical Services ambulances, and 957 of them were infants younger than one year (24.2%). The fact that infants under one year accounted for nearly one quarter of ambulance admissions may be related to the increased clinical vulnerability of this age group and their greater need for urgent medical evaluation. Among the included infants, 532 (55.6%) were male, and 425 (44.4%) were female (Table 1). The mean birth weight was 2,957.4 g, while the mean maternal age was 28.4 years.

**Table 1.** Distribution of gender and admission reasons among infants aged 0–12 months transported to the pediatric emergency department by 112 emergency medical services ambulances.

| Variable | Groups | n | % |
|---|---|---|---|
| Gender | Male | 532 | 55.6 |
| | Female | 425 | 44.4 |
| | Total | 957 | 100.0 |
| **Reason for admission** | | **n** | **%** |
| Respiratory distress | | 371 | 38.8 |
| Jaundice | | 209 | 21.8 |
| Fever | | 79 | 8.3 |
| Feeding difficulties | | 67 | 7.0 |
| Vomiting | | 56 | 5.9 |
| Trauma | | 34 | 3.6 |
| Constipation | | 26 | 2.7 |
| Dermatological conditions | | 22 | 2.3 |
| Diarrhea | | 14 | 1.5 |
| Irritability and agitation | | 19 | 2.0 |
| Cough | | 16 | 1.7 |
| Surgical conditions | | 12 | 1.3 |
| Cystitis | | 11 | 1.1 |
| Neurological conditions | | 3 | 0.3 |
| Other | | 18 | 1.9 |
| Total | | 957 | 100.0 |

When the reasons for admission to the pediatric emergency department via 112 Emergency Medical Services ambulances were evaluated, respiratory distress was the most frequent complaint, accounting for 371 cases (38.8%). Jaundice followed as the second most common reason with 209 cases (21.8%), while fever ranked third with 79 cases (8.3%). Other presenting complaints included feeding difficulties, vomiting, trauma, constipation, dermatological problems, diarrhea, irritability and restlessness, cough, surgical conditions, cystitis, and neurological disorders (Table 1). The "other" category comprised conditions such as thyroid function abnormalities (elevated thyroid-stimulating hormone levels), sleep-related problems, conjunctivitis, prematurity, burns, and hypothermia.

Of the patients transported by 112 Emergency Medical Services ambulances, 580 (60.6%) were hospitalized. Among them, 116 (20.0%) required admission to the Neonatal Intensive Care Unit, most commonly due to respiratory distress, jaundice, and fever. A total of 356 patients (37.2%) who did not require hospitalization were discharged following clinical assessment and appropriate management. Additionally, 21 patients (2.2%) were referred to other centers for further diagnostic evaluation and treatment. The main reasons for referral were trauma or fall history (n=12), fever (n=5), severe respiratory distress (n=3), and jaundice (n=1).

When repeat visits to the pediatric emergency department were analyzed, 108 patients (11.3%) had at least one readmission during the study year via 112 Emergency Medical Services ambulances. Respiratory distress was the leading cause of these repeat admissions (59.3%), followed by jaundice (25.9%) and fever (14.8%).

The diagnostic tests and imaging methods used for patients evaluated in the pediatric emergency department are presented in Table 2. Laboratory blood tests were the most commonly requested investigations in this cohort. No diagnostic testing was performed in 105 patients (11.1%), while 566 patients (59.1%) underwent laboratory blood analysis. Among other diagnostic and imaging modalities, ultrasonography was used in 96 patients (10.0%), plain radiography in 67 patients (7.0%), urine analysis in 54 patients (5.6%), computed tomography in 43 patients (4.5%), and magnetic resonance imaging in 26 patients (2.7%).

When the timing of emergency department presentations among infants aged 0–12 months was examined, 283 admissions (29.6%) occurred on weekdays, while 674 (70.4%) took place on weekends and official holidays (Table 3). In terms of time of day, 108 presentations (11.3%) were recorded between 08:00 and 16:00, 617 (64.5%) between 16:00 and 24:00, and 232 (24.2%) between 24:00 and 08:00 (Table 3).

**Table 2.** Diagnostic tests and imaging studies requested for infants aged 0–12 months transported to the pediatric emergency department by 112 emergency medical services ambulances.

| Diagnostic test information | n | % |
|---|---|---|
| Laboratory blood tests requested | 566 | 59.1 |
| No diagnostic tests requested | 105 | 11.1 |
| Ultrasonography (USG) requested | 96 | 10.0 |
| Plain radiography requested | 67 | 7.0 |
| Urine tests requested | 54 | 5.6 |
| Computed tomography (CT) requested | 43 | 4.5 |
| Magnetic resonance imaging (MRI) requested | 26 | 2.7 |
| Total | 957 | 100.0 |

**Table 3.** Timing and distribution of admissions of infants aged 0–12 months transported to the pediatric emergency department by 112 emergency medical services ambulances.

| Time of admission | n | % |
|---|---|---|
| Weekday admissions | 283 | 29.6 |
| Weekend and public holiday admissions | 674 | 70.4 |
| Total | 957 | 100.0 |
| **Time of admission** | **n** | **%** |
| Admissions between 08:00 and 16:00 | 108 | 11.3 |
| Admissions between 16:00 and 24:00 | 617 | 64.5 |
| Admissions between 24:00 and 08:00 | 232 | 24.2 |
| Total | 957 | 100.0 |

## DISCUSSION

In our study, analysis of the reasons for admission to the pediatric emergency department via 112 Emergency Medical Services ambulances showed that respiratory distress was the most frequent presenting complaint, with jaundice ranking second. A study conducted at the Çukurova University Faculty of Medicine Hospital similarly identified respiratory distress as the leading cause of emergency visits among preterm infants, which aligns with our findings[6]. In addition, Bozlu et al. reported that the most common reasons for neonatal admissions to the pediatric emergency department were jaundice, prematurity, and asphyxia, and that 78.4% of these patients required hospitalization[7].

According to our findings, the mean birth weight of infants presenting to the pediatric emergency department was 2,957.4 g. In a similar study, the mean birth weight was reported as 3,232.22±456 g[8]. Among the infants included in our study, 532 (55.6%) were male and 425 (44.4%) were female, indicating a higher proportion of male patients. This finding is in line with previous studies evaluating pediatric emergency department admissions, which have also reported a higher frequency of male patients compared to females[9].

In our study, the most common indications for admission to the neonatal intensive care unit were respiratory distress and jaundice. In contrast, another study in the literature reported that infants presenting to the emergency department with complaints such as fever and dehydration had higher hospitalization rates compared with other conditions. Based on these findings, it has been suggested that infants presenting with these clinical features should be admitted promptly for observation and treatment, without unnecessary delay in the emergency department for further diagnostic evaluation, provided their general condition is stable[10].

As observed in this study, admissions of patients transported to the pediatric emergency department by 112 Emergency Medical Services ambulances occurred more frequently on weekends and public holidays than on weekdays. This pattern is consistent with previous literature describing the temporal distribution of pediatric emergency department visits. In a large-scale observational study by Schinkelshoek et al., a considerable proportion of pediatric emergency presentations were reported during out-of-office hours and on weekends, with higher patient volumes during these time periods[11]. Limited access to primary healthcare services outside regular working hours, increased parental concern when symptoms arise, and the inherent vulnerability of the neonatal period may all contribute to this pattern. In particular, symptoms such as respiratory distress, fever, and feeding difficulties in neonates often lead to heightened parental anxiety, which in turn increases the use of emergency medical services.

When the findings of this study are considered together with the existing literature, a considerable proportion of pediatric emergency department admissions via 112 Emergency Medical Services ambulances appears to occur during the early neonatal period. Previous studies have shown that emergency department visits among neonates are most commonly clustered within the first week of life[12]. This pattern reflects the physiological vulnerability of this early stage and the increased need for urgent medical care. The results emphasize the importance of providing comprehensive postnatal parental education and maintaining close follow-up for high-risk neonates during the early postnatal period.

## CONCLUSION

Infants represent a considerable proportion of pediatric emergency department admissions due to being in a highly vulnerable stage of life. In this study, a notable proportion of patients (37.2%) transported to the pediatric emergency department by 112 Emergency Medical Services ambulances were discharged following physical examination and basic clinical evaluation alone. This finding indicates that many emergency department visits may not correspond to true emergency conditions. In most cases, clinical assessment and routine laboratory investigations were sufficient for diagnosis and management, whereas the need for advanced imaging techniques and invasive procedures was limited.

The findings of this study suggest that insufficient parental education during the early postnatal period regarding physiological changes, routine well-child follow-up, and warning signs may be an important factor contributing to increased emergency department visits. In addition, early discharge practices and limited outpatient follow-up may also contribute to the rising number of pediatric emergency department admissions among infants aged 0–12 months. Strengthening postnatal parental education, establishing structured and regular outpatient follow-up for high-risk infants, and encouraging more appropriate use of primary healthcare services may help reduce unnecessary emergency department visits and repeat admissions.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Mardin Training and Research Hospital Scientific Research Ethics Committee (approval date: October 8, 2024; approval number: 2024/10-27). The requirement for informed consent was waived due to the retrospective nature of the study.

## AUTHORS' CONTRIBUTIONS

**SC:** Conceptualization. **MKN:** Data curation, Writing – original draft. **ED:** Data curation, Writing – review & editing. **MS:** Writing – review & editing. **MEP:** Formal Analysis. **AA:** Formal Analysis.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Paul DA, Agiro A, Hoffman M, Denemark C, Brazen A, Pollack M, et al. Hospital admission and emergency department utilization in an infant medicaid population. Hosp Pediatr. 2016;6(10):587-94. https://doi.org/10.1542/hpeds.2015-0254
2. Morgans A, Burgess SJ. What is a health emergency? The difference in definition and understanding between patients and health professionals. Aust Health Rev. 2011;35(3):284-9. https://doi.org/10.1071/AH10922
3. Sammour I, Karnati S, Othman H, Heis F, Peluso A, Aly H. Trends in procedures in the neonatal intensive care unit. Am J Perinatol. 2024;41(S 01):e494-500. https://doi.org/10.1055/a-1905-5245
4. Wang XL, Wang J, Yuan L, Shi WJ, Cao Y, Chen C. Trend and causes of neonatal mortality in a level III children's hospital in Shanghai: a 15-year retrospective study. World J Pediatr. 2018;14(1):44-51. https://doi.org/10.1007/s12519-017-0101-y
5. Alpern ER, Stanley RM, Gorelick MH, Donaldson A, Knight S, Teach SJ, et al. Epidemiology of a pediatric emergency medicine research network: the PECARN Core Data Project. Pediatr Emerg Care. 2006;22(10):689-99. https://doi.org/10.1097/01.pec.0000236830.39194.c0
6. Atici A, Pirti M, Turkmen M, Narli N, Satar M. Problems of neonates admitted to the emergency department. Turkiye Klinikleri J Pediatr. 1996;5(2):49-52.
7. Bozlu G, Sagcan MF, Celik Y, Kuyucu N. Retrospective evaluation of 624 neonates admitted to the pediatric emergency department. Mersin Univ Saglik Bilim Derg. 2018;11(1):7-12.
8. Simsek Pervane S, Cicek C, Halil H. Evaluation of neonatal profile presenting to the pediatric emergency department. Anatolian J Emerg Med. 2024;7(1):32-6. https://doi.org/10.54996/anatolianjem.1396395
9. Rhinehart B, Sheridan D, Chung S, McGaughey S. Neonatal presentations in the pediatric emergency department: a decade-long retrospective analysis. Am J Emerg Med. 2025;87:65-9. https://doi.org/10.1016/j.ajem.2024.10.041
10. Gunes S, Sahin S, Koyuncu Arslan M, Karaca Dag O, Anil M. Emergency room admission for newborns: how many are really urgent? Forbes J Med. 2022;3(3):291-96. https://doi.org/10.4274/forbes.galenos.2022.73644
11. Schinkelshoek G, Borensztajn DM, Zachariasse JM, Maconochie IK, Alves CF, Freitas P, et al. Management of children visiting the emergency department during out-of-office hours: an observational study. BMJ Paediatr Open. 2020;4(1):e000687. https://doi.org/10.1136/bmjpo-2020-000687
12. Millar KR, Gloor JE, Wellington N, Joubert GI. Early neonatal presentations to the pediatric emergency department. Pediatr Emerg Care. 2000;16(3):145-50. https://doi.org/10.1097/00006565-200006000-00001

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260594

# The role of hematological parameters and systemic immune-inflammation index in the management of symptomatic gestational hydronephrosis

Ubeyd Sungur[1]*, Kağan Toksal[1], Gül Şüheda Sungur[2], Deniz Noyan Özlü[1], Mithat Ekşi[1], Ali Emre Fakir[1], Alper Bitkin[1]

## SUMMARY

**OBJECTIVE:** The aim of this study was to investigate the clinical, imaging, and hematological parameters predicting the necessity of surgical intervention in pregnant women presenting with symptomatic gestational hydronephrosis.

**METHODS:** We retrospectively evaluated 192 pregnant women diagnosed with symptomatic hydronephrosis between January 2015 and September 2024. Patients were divided into Group 1 (conservative management, n=138) and Group 2 (surgical intervention via double-J stent or percutaneous nephrostomy, n=54). Demographic data, clinical findings, degree of hydronephrosis, urinalysis, and complete blood count parameters—including C-reactive protein, neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, and systemic immune-inflammation index—were compared. Logistic regression analysis was performed to identify predictors for surgery.

**RESULTS:** A history of urolithiasis, prolonged pain, fever, severe hydronephrosis, pyuria, and elevated creatinine were significantly higher in Group 2 ($p<0.05$). White blood cell, neutrophil-to-lymphocyte ratio, and C-reactive protein levels did not differ significantly between the groups. However, platelet counts ($p=0.020$), platelet-to-lymphocyte ratio ($p=0.038$), and systemic immune-inflammation index ($p=0.015$) were significantly elevated in the surgical cohort. Univariate logistic regression analysis revealed that an elevated platelet count ($p=0.039$) and a higher systemic immune-inflammation index ($p=0.004$) were significant predictors. In the multivariate analysis, the systemic immune-inflammation index remained an independent predictor ($p=0.009$), whereas platelet count did not.

**CONCLUSION:** While clinical presentation and the severity of hydronephrosis primarily indicate the need for surgical intervention, the systemic immune-inflammation index serves as a valuable, cost-effective auxiliary predictor among the hematological parameters.

**KEYWORDS:** Hydronephrosis. Pregnancy complications. Ureteral obstruction. Inflammation. Biomarkers.

## INTRODUCTION

Gestational hydronephrosis generally presents as a transient condition that does not require invasive treatment. It is well established that this condition is caused by the mechanical compression of the growing fetus on the ureter—predominantly on the right side—coupled with the relaxant effect of progesterone[1]. While it is frequently detected via ultrasonography during routine check-ups in asymptomatic pregnant women, it occasionally presents with flank pain. Given the restricted imaging modalities during pregnancy, the possibility of ureteral stones must always be kept in mind[2]. Particular attention is paid to identify the management method, as interventions performed during pregnancy or delayed treatments can adversely affect both the mother and the fetus[3]. Currently, the criteria for determining the treatment modality in symptomatic gestational hydronephrosis lack clearly defined boundaries.

A known history of stone disease and recent prior imaging, if available, can be important in the decision-making process for surgical intervention in this patient cohort. In cases presenting with signs of pyelonephritis, fever, pyuria, prolonged intractable pain, or renal impairment, surgical alternatives such as double-J (DJ) stent placement or percutaneous nephrostomy are considered[4]. A risk-benefit analysis is conducted in collaboration with the patient, taking the remaining time until delivery into account. Although clinical findings primarily constitute the determining factors, the predictive value of inflammatory markers in this patient group has become a subject of ongoing research[5,6]. While mild leukocytosis is considered a normal physiological change during pregnancy, the presence of severe

[1]University of Health Sciences, Bakırköy Dr. Sadi Konuk Research and Training Hospital, Department of Urology – Istanbul, Turkey.
[2]University of Health Sciences, Bakırköy Dr. Sadi Konuk Research and Training Hospital, Department of Obstetrics and Gynecology – Istanbul, Turkey.
*Corresponding author: ubeydsungur@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on April 06, 2026. Accepted on July 19, 2026.
Scientific Editor: José Maria Soares Júnior

leukocytosis and neutrophilia, along with elevated parameters such as C-reactive protein (CRP), may indicate an impending septic state and thus aid in clinical decision-making[7,8]. Furthermore, parameters including the neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and systemic immune-inflammation index (SII) have emerged as hematological markers widely investigated for their role in signifying inflammation across various diseases[9,10].

In this study, we aimed to investigate the hematological markers that predict the necessity for surgical intervention by comparing the clinical, imaging, and hematological parameters of pregnant women evaluated for symptomatic hydronephrosis who either underwent surgical intervention or were managed with conservative treatment.

## METHODS

Pregnant women diagnosed with hydronephrosis via ultrasonography during routine obstetric follow-ups or presentations to the emergency department, and who were subsequently referred to the urology department at our hospital between January 2015 and September 2024, were retrospectively reviewed. Based on patient history and consultation notes, symptomatic patients presenting with complaints such as pain, fever, or dysuria were included in the study. Patients with a solitary kidney, a history of renal transplantation, an immunodeficiency disorder, or missing essential laboratory data were excluded.

Upon admission, all patients underwent a complete blood count (CBC), biochemical evaluation including serum creatinine and CRP, urinalysis, and urinary ultrasonography. Patient histories were noted for gestational trimester, parity, presence of fever, duration of pain, and a history of urolithiasis. The patients were classified into two groups for comparison: Group 1 consisted of patients managed with conservative treatment, and Group 2 comprised those who underwent surgical intervention. Conservative management included fluid hydration, analgesic support, and antibiotherapy. Based on our institutional clinical practice, the indications for surgical intervention were defined as follows: (1) signs of pyelonephritis, defined as the coexistence of fever, pyuria, and hydronephrosis of any grade; (2) severe flank pain persisting for more than three days accompanied by hydronephrosis of any grade; or (3) severe hydronephrosis accompanied by elevated creatinine and/or CRP levels. Surgical interventions involved the placement of a 4.7-French (Fr) DJ stent under general anesthesia or sedation, or the insertion of an 8-Fr percutaneous nephrostomy tube under local anesthesia. The choice between the two procedures was determined by the surgeon according to the severity of hydronephrosis and patient preference. In the urinalysis, the presence of ≥10 leukocytes per high-power field (HPF) was defined as pyuria, and the presence of ≥4 erythrocytes per HPF was defined as hematuria.

The two groups were compared regarding demographic characteristics, clinical findings, the degree of hydronephrosis, and laboratory values. Subsequently, hematological parameters were subjected to logistic regression analysis to identify predictors for the necessity of surgical intervention.

### Statistical analysis

Categorical data were presented as frequencies and percentages. The Kolmogorov-Smirnov test was utilized to evaluate the normal distribution of numerical data. Normally distributed continuous variables were expressed as mean±standard deviation, whereas non-normally distributed continuous variables were reported as median and interquartile range (IQR). The Student's t-test was used to compare normally distributed numerical data, and the Mann-Whitney U test was applied for non-normally distributed data. The frequencies of categorical variables were compared using Pearson's chi-square test and Fisher's exact test. Univariate logistic regression analysis was performed to determine the predictive value of hematological parameters for surgical intervention. Parameters found to be significant in the univariate analysis (platelet count and SII) were evaluated for multicollinearity by calculating the variance inflation factor (VIF) and were subsequently entered into a multivariate logistic regression model. Receiver operating characteristic (ROC) curve analysis was then performed to assess the discriminative ability of the SII, and the area under the curve (AUC) with its 95%CI, the optimal cut-off value (determined by the Youden index), sensitivity, and specificity were calculated. A p-value of <0.05 was considered statistically significant. All statistical analyses were performed using the Statistical Package for the Social Sciences version 26.0 (IBM SPSS Statistics; IBM Corp., Armonk, NY, USA).

## RESULTS

In our study evaluating 192 pregnant women, the patients were divided into two groups: those managed with conservative treatment (Group 1, n=138) and those who underwent surgical treatment (Group 2, n=54). Within Group 2, DJ stent placement was performed in 36 patients (66.7%), while 18 patients (33.3%) underwent percutaneous nephrostomy insertion. The mean age of the total cohort was 26.6±5.1 years. Age, trimester of pregnancy, and primiparity rates did not differ significantly between the two groups (p>0.05). A known history

of urolithiasis, the duration of pain before admission, and the presence of fever were found to be significantly higher in Group 2 (p=0.011, p=0.001, and p<0.001, respectively). While there was no significant difference between the groups regarding the laterality of hydronephrosis, the rate of severe hydronephrosis was considerably higher in Group 2 (p<0.001).

Regarding the CBC parameters, no significant differences were detected between the two groups in terms of white blood cell (WBC), neutrophil, and lymphocyte counts. However, the platelet count was notably higher in Group 2 compared to Group 1 (p=0.02). The median values of the NLR and CRP were higher in Group 2 than in Group 1; however, these differences did not reach statistical significance (p=0.154 and p=0.248, respectively). Conversely, PLR and SII values were remarkably higher in Group 2 in contrast to Group 1 (p=0.038 and p=0.015, respectively). Elevated creatinine levels and the presence of pyuria in urinalysis were significantly higher in Group 2, whereas no significant difference was observed between the groups concerning hematuria. The detailed comparison between the two groups is presented in Table 1.

According to the univariate logistic regression analysis conducted to investigate the hematological predictors for the necessity of surgical intervention, elevated platelet count and SII values (calculated by dividing by 100) were found to be statistically significant predictors (p=0.039 and p=0.004, respectively). Prior to the multivariate analysis, these two significant parameters were assessed for multicollinearity; the correlation between them was weak (r=0.20), and the variance inflation

**Table 1.** Comparison of demographic, clinical, and laboratory parameters of the patients between the conservative management and surgical treatment groups.

| Parameters (mean±SD; %) | Total (n=192) | Group 1 "Conservative" n=138 (71.9) | Group 2 "Surgical" n=54 (28.1) | p-value |
|---|---|---|---|---|
| Age (years) | 26.6±5.1 | 26.5±5.4 | 27±4.6 | 0.585 |
| Trimester (n; %) | | | | |
| First trimester | 19 (9.9) | 10 (7.2) | 9 (16.7) | 0.136 |
| Second trimester | 96 (50) | 70 (50.8) | 26 (48.1) | |
| Third trimester | 77 (40.1) | 58 (42) | 19 (35.2) | |
| Primiparous (n; %) | 114 (59.4) | 81 (58.7) | 33 (61.1) | 0.759 |
| History of urolithiasis (n; %) | 41 (21.4) | 23 (16.7) | 18 (33.3) | **0.011** |
| Duration of flank pain (days), median (IQR) | 3 (4) | 2 (3) | 3 (5) | **0.001** |
| Fever (n; %) | 29 (15.1) | 12 (8.7) | 17 (31.5) | **<0.001** |
| Hydronephrosis (n; %) | | | | |
| Grade 1–2 | 138 (71.9) | 116 (84.1) | 22 (40.7) | **<0.001** |
| Grade 3–4 | 54 (28.1) | 22 (15.9) | 32 (59.3) | |
| Laterality of hydronephrosis (n; %) | | | | |
| Right | 118 (61.5) | 87 (63) | 31 (57.4) | 0.651 |
| Left | 35 (18.2) | 23 (16.7) | 12 (22.2) | |
| Bilateral | 39 (20.3) | 28 (20.3) | 11 (20.4) | |
| White blood cells ($\times10^3/\mu L$) | 13.6±4.3 | 13.2±4.2 | 14.5±4.3 | 0.074 |
| Neutrophil count ($\times10^3/\mu L$) | 10.9±4.1 | 10.6±3.9 | 11.7±4.5 | 0.106 |
| Lymphocyte count ($\times10^3/\mu L$) | 1.69±0.66 | 1.71±0.65 | 1.66±0.69 | 0.647 |
| Platelet count ($\times10^3/\mu L$), median (IQR) | 244 (78) | 237 (77.7) | 260.5 (68.5) | **0.020** |
| NLR, median (IQR) | 6.2 (5.76) | 5.88 (5.22) | 7.16 (6.86) | 0.154 |
| PLR, median (IQR) | 152.8 (86.8) | 147.8 (84) | 169.8 (95.9) | **0.038** |
| SII, median (IQR) | 1,593 (1,625) | 1,465 (1,259) | 1,817 (2,141) | **0.015** |
| CRP (mg/L), median (IQR) | 12 (33) | 11 (43.7) | 16.5 (27.7) | 0.248 |
| Creatinine (mg/dL), median (IQR) | 0.5 (0.21) | 0.49 (0.18) | 0.56 (0.33) | **0.002** |
| Pyuria (n; %) | 91 (47.4) | 56 (40.6) | 35 (64.8) | **0.002** |
| Hematuria (n; %) | 160 (83.3) | 112 (81.2) | 48 (88.9) | 0.196 |

SD: standard deviation; NLR: neutrophile to lymphocyte ratio; PLR: platelet to lymphocyte ratio; IQR: interquartile range; SII: systemic immune-inflammation index; CRP: C-reactive protein. Bold values indicate statistical significance (p<0.05).

**Table 2.** Univariate and multivariate logistic regression analysis of hematological parameters for predicting the need for surgical intervention.

| | Univariate Analysis | | | Multivariate Analysis | | |
|---|---|---|---|---|---|---|
| | OR | 95%CI | p | OR | 95%CI | p |
| WBC | 1.067 | 0.993–1.146 | 0.078 | | | |
| Neutrophil | 1.063 | 0.986–1.146 | 0.109 | | | |
| Lymphocyte | 0.894 | 0.554–1.441 | 0.645 | | | |
| Platelet | 1.004 | 1.000–1.008 | **0.039** | 1.003 | 0.999–1.008 | 0.125 |
| NLR | 1.029 | 0.981–1.078 | 0.244 | | | |
| PLR | 1.004 | 1.000–1.008 | 0.070 | | | |
| SII (per 100 units) | 1.036 | 1.012–1.062 | **0.004** | 1.033 | 1.008–1.058 | **0.009** |
| CRP | 1.002 | 0.998–1.007 | 0.321 | | | |

OR: odds ratio; CI: confidence interval; WBC: white blood cell; NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; SII: systemic immune-inflammation index; CRP: C-reactive protein. Bold values indicate statistical significance (p<0.05).

factors were low (1.09 for platelet count and 1.35 for SII), indicating no significant multicollinearity. When both parameters were entered into the multivariate logistic regression model, the SII remained an independent predictor of the need for surgical intervention (OR 1.033, 95%CI 1.008–1.058, p=0.009), whereas the platelet count lost its statistical significance (OR 1.003, 95%CI 0.999–1.008, p=0.125). The results of the logistic regression analyses are outlined in Table 2. In the ROC curve analysis, the SII showed a limited discriminative ability for predicting the need for surgical intervention, with an area under the curve (AUC) of 0.613 (95%CI 0.517–0.705). The optimal cut-off value was 2,865, yielding a sensitivity of 40.7% and a specificity of 87.7%. Figure 1 illustrates the graphical comparison of SII between the two groups. Figure 2 illustrates the ROC curve of the SII for predicting the need for surgical intervention.

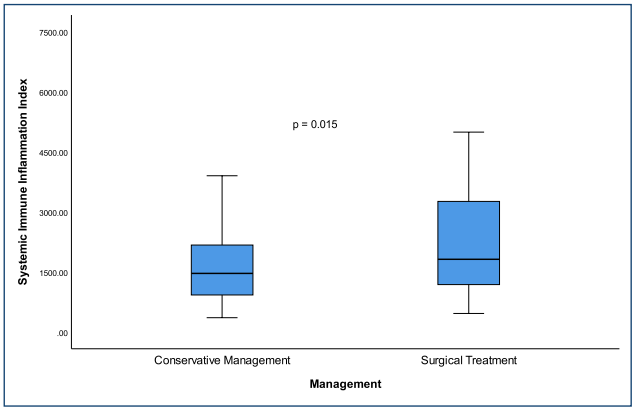



**Figure 1.** Comparison of the systemic immune-inflammation index between the conservative management group (Group 1) and the surgical treatment group (Group 2) (Mann-Whitney U test, p=0.015).

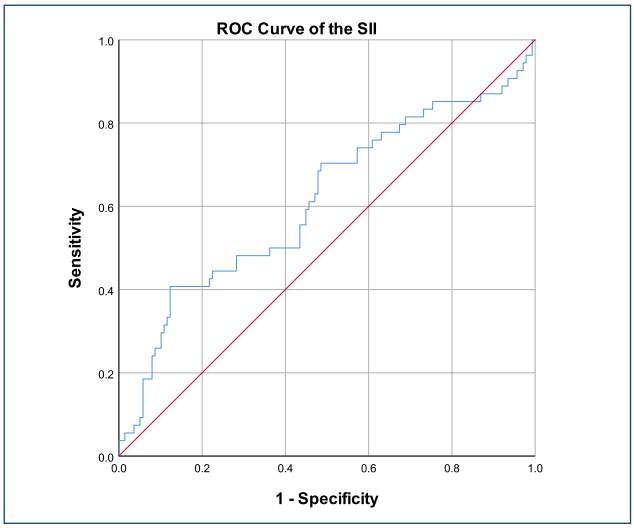



**Figure 2.** Receiver operating characteristic curve of the systemic immune-inflammation index for predicting the need for surgical intervention.

## DISCUSSION

Our study investigated the relationship between hematological parameters and the necessity for surgical intervention over 9 years by comparing pregnant women presenting with symptomatic hydronephrosis at our center who underwent either surgical procedures or conservative management. While clinical findings such as fever, prolonged pain, pyuria, and severe hydronephrosis on ultrasonography were observed at significantly higher rates in patients who underwent surgical intervention, the SII emerged as the most robust predictor among the evaluated blood parameters for the necessity of surgery.

During management of symptomatic hydronephrosis in pregnancy, the impact of inflammation on maternal health and the risk of infection must be carefully reviewed against the potential effects of interventional procedures on the fetus

and overall obstetric outcomes[11]. Analyzing our results, we observed that clinical findings were the primary determining factors in patients who underwent percutaneous nephrostomy or DJ stent placement. Drainage is typically planned when findings of pyelonephritis, such as fever, pyuria, and flank pain, are accompanied by hydronephrosis indicative of urinary obstruction. Since computed tomography is contraindicated in this patient population, decision-making relies heavily on ultrasonography and clinical presentation. Consistent with this, our study found significantly higher rates of fever, prolonged pain, pyuria, and elevated creatinine levels in patients requiring surgical intervention. The elevated creatinine levels in the surgical group likely reflect the functional impact of urinary obstruction on renal drainage. Because glomerular filtration physiologically increases during pregnancy and lowers baseline serum creatinine, even values within or near the upper normal range may indicate subclinical renal impairment secondary to obstruction; a rising creatinine in this setting may therefore support the decision for decompression to preserve renal function. In a study examining pregnant women presenting with renal colic, He et al. reported that pain lasting longer than 4 days was a predictor for surgical intervention[12]. The severity of hydronephrosis also influences the clinician's decision-making process by helping to differentiate pathological obstruction from physiological gestational hydronephrosis. The degree of hydronephrosis and anteroposterior pelvic diameter have been reported in the literature as parameters predicting the need for surgical intervention in pregnant women with hydronephrosis[13]. Our findings strongly support these existing data. Furthermore, it has been established that the presence of urolithiasis during pregnancy poses significant risks to both maternal and neonatal health[14]. Considering the recurrent nature of stone disease, we believe that in pregnant women with a history of urolithiasis who develop hydronephrosis, clinicians may be more inclined toward surgical intervention even if a stone is not directly visualized.

There are several studies in the literature evaluating the use of blood parameters in the decision-making process for symptomatic gestational hydronephrosis. Ercil et al. stated that elevated WBC counts and CRP levels, in conjunction with pain, are markers for the necessity of surgical intervention[5]. Conversely, a study published by Demir et al. found no significant difference in WBC and CRP elevations between patient groups managed surgically and those followed conservatively[15]. In our study, although WBC and CRP were higher in the surgical group, the difference did not reach statistical significance. It is well known that pregnancy-induced leukocytosis can occur, meaning WBC may not always accurately reflect an underlying pathological state[16]. Considering that it could be helpful without requiring additional laboratory testing, Bozkurt et al. investigated the NLR and identified that an elevated NLR could serve as an auxiliary parameter in the decision for surgical intervention[6]. Although the NLR was observed to be higher in the surgical group of our study, we did not detect a statistically significant difference. The SII, calculated by multiplying the NLR by the platelet count, has emerged as another valuable inflammatory marker. Previous literature has associated the SII with certain obstetric pathologies, including preeclampsia, early gestational diabetes, and threatened abortion[17-19]. In our cohort, the SII was found to be significantly elevated in patients undergoing surgery due to hydronephrosis and served as a significant predictor for the necessity of surgery. It is noteworthy that, while the individual leukocyte-based markers (WBC, CRP, and NLR) did not reach statistical significance, the composite SII did. This apparent paradox may be attributed to the fact that the SII additionally incorporates the platelet-driven component of the inflammatory response, which is not captured by the leukocyte-based indices alone, whereas the physiological leukocytosis of pregnancy may further attenuate the discriminative value of the WBC and NLR. Notably, although the PLR reached significance in the baseline group comparison (p=0.038), it did not retain significance in the univariate regression model (p=0.070), further underscoring the relative stability of the SII. Consistent with this, in the multivariate model the SII remained an independent predictor, whereas the platelet count alone did not, indicating that it is the combination of these components rather than the platelet count in isolation that carries the predictive signal. However, its limited discriminative ability on ROC analysis (AUC=0.613) indicates that the SII should be regarded as an adjunct to clinical and radiological assessment rather than a standalone predictor. It should also be emphasized that, although statistically significant, the odds ratios of these hematological parameters reflect modest per-unit increments (for instance, an OR of 1.033 per 100-unit increase in the SII). Accordingly, the value of the SII lies in its ability to add an accessible, objective, and quantifiable dimension to an otherwise clinically driven decision, complementing rather than competing with established clinical and radiological parameters.

In a normal, uncomplicated pregnancy, the platelet count may decrease slightly[20]. However, in the presence of inflammation, an increase in platelet count may be observed due to the activation of platelets and their role in the transport of cytokines[21]. In gestational hydronephrosis, the platelet count may be elevated due to inflammation secondary to obstruction, which consequently leads to higher calculated PLR and SII values. While the decision for

surgical intervention in gestational hydronephrosis is primarily driven by clinical findings and imaging results, we believe that incorporating parameters such as SII into the evaluation process will provide valuable diagnostic assistance.

Our study has certain limitations. These include its retrospective design, single-center nature, and relatively small sample size. In addition, the multivariable model was restricted to hematological parameters and was not adjusted for clinical variables. However, it is particularly significant as it is the first study to investigate the relationship between the SII and symptomatic gestational hydronephrosis. We believe that multicenter prospective studies with larger patient cohorts are required to validate these findings further.

## CONCLUSION

The management of symptomatic gestational hydronephrosis relies predominantly on clinical presentation and ultrasound findings, with severe hydronephrosis, prolonged pain, fever, and pyuria strongly indicating the need for surgical intervention. However, our study demonstrates that the SII, an easily accessible and cost-effective marker, can serve as a valuable auxiliary predictor among the hematological parameters, while its limited discriminative ability indicates that it should complement rather than replace clinical and radiological assessment. As the first study to highlight the predictive value of the SII in this specific patient cohort, further studies are warranted to validate these findings.

## ETHICS COMMITTEE APPROVAL

The study was approved by the University of Health Sciences, Dr. Sadi Konuk Training and Research Hospital Ethical Committee, Bakirkoy, Istanbul, Turkey (Approval Number: 2025/381, Decision Date: 07.11.2025).

## INFORMED CONSENT

An informed consent was obtained from all the patients. This study was approved by the ethics committee of University of Health Sciences, Dr. Sadi Konuk Training and Research Hospital Ethical Committee (Approval Number: 2025/381, Decision Date: 07.11.2025).

## AUTHORS' CONTRIBUTIONS

**US:** Conceptualization, Data curation, Writing – original draft. **KT:** Data curation. **GŞS:** Conceptualization, Writing – review & editing. **DNÖ:** Writing – review & editing. **ME:** Conceptualization, Writing – review & editing. **AEF:** Writing – review & editing. **AB:** Supervision.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Demiroğlu Ç, Solakhan M. Hydronephrosis in pregnancy: critical factors determining urinary catheter use. Obstet Med. 2026;19:40-5.
2. Keenan RA, Hegarty NJ, Davis NF. Symptomatic hydronephrosis and ureteral calculi in pregnancy: a narrative review with a proposed management protocol. J Endourol. 2022;36:1099-112.
3. Hosny M, Chan K, Ibrahim M, Sharma V, Vasdev N. The management of symptomatic hydronephrosis in pregnancy. Cureus. 2024;16:e52146.
4. Tsai YL, Seow KM, Yieh CH, Chong KM, Hwang JL, Lin YH, et al. Comparative study of conservative and surgical management for symptomatic moderate and severe hydronephrosis in pregnancy: a prospective randomized study. Acta Obstet Gynecol Scand. 2007;86:1047-50.
5. Ercil H, Arslan B, Ortoglu F, Alma E, Unal U, Deniz ME, et al. Conservative/surgical treatment predictors of maternal hydronephrosis: results of a single-center retrospective non-randomized non-controlled observational study. Int Urol Nephrol. 2017;49:1347-52.
6. Bozkurt M, Seker KG, Erkoc M, Danis E, Can O, Degirmentepe RB, et al. Value of neutrophil-to-lymphocyte ratio in predicting surgical intervention for hydronephrosis during pregnancy. Urologia. 2023;90:365-70.
7. Conwell J, Ayyash M, Singh HK, Goffman D, Ranard BL. Physiologic changes of pregnancy and considerations for screening and diagnosis of sepsis. Semin Perinatol. 2024;48:151973.
8. Dockree S, Brook J, James T, Shine B, Impey L, Vatish M. Pregnancy-specific reference intervals for C-reactive protein improve diagnostic accuracy for infection: a longitudinal study. Clin Chim Acta. 2021;517:81-5.
9. Martinez JM, Espírito Santo A, Ramada D, Fontes F, Medeiros R. Diagnostic accuracy of neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, and neutrophil-lymphocyte-to-platelet ratio biomarkers in predicting bacteremia and sepsis in immunosuppressive patients with cancer: literature review. Porto Biomed J. 2024;9:254.
10. Çallıoğlu N, Tuna G, Tandoğan Ö, Ersan F, Atalay S, Bilirer KK. Systemic immune-inflammatory index and platelet-to-lymphocyte ratio in intrahepatic cholestasis of pregnancy. Saudi Med J. 2024;45:1217-22.
11. Fainaru O, Almog B, Gamzu R, Lessing JB, Kupferminc M. The management of symptomatic hydronephrosis in pregnancy. BJOG. 2002;109:1385-7.
12. He M, Lin X, Lei M, Xu X, He Z. The identification of pregnant women with renal colic who may need surgical intervention. BMC Urol. 2022;22:30.

13. Haberal HB, Tonyali S. Factors predicting the need for surgical intervention for hydronephrosis during pregnancy: a systematic review of the literature. Arch Gynecol Obstet. 2024;309:1801-6.
14. Sebastian N, Czuzoj-Shulman N, Spence AR, Abenhaim HA. Maternal and fetal outcomes of urolithiasis: a retrospective cohort study. J Gynecol Obstet Hum Reprod. 2021;50:102161.
15. Demir M, Yağmur İ, Pelit ES, Katı B, Tunçtekin A, Çiftçi H. Comparison of conservative and surgical treatments in symptomatic pregnancy hydronephrosis. Urol Int. 2021;105:1085-91.
16. Bridwell RE, Carius BM, Long B, Oliver JJ, Schmitz G. Sepsis in pregnancy: recognition and resuscitation. West J Emerg Med. 2019;20:822-32.
17. Zhang T, He X. Systemic immune-inflammation index and risk of gestational diabetes and preeclampsia: a systematic review and meta-analysis. J Obstet Gynaecol. 2025;45:2548814.
18. Li Q, Li M, Zhao C, Chen X, Zhang C, Li Y, et al. Association between systemic immune-inflammation index in early pregnancy and preeclampsia: a multicenter cohort study. J Inflamm Res. 2025;18:8107-18. https://doi.org/10.2147/JIR.S523131
19. Soysal C, Sarı H, Işıkalan MM, Özkaya EB, Ulaş Ö, Taşçı Y, et al. Role of the systemic immune-inflammation index in threatened abortion patients and predicting of abortion. J Obstet Gynaecol Res. 2023;49:1723-8.
20. Larroca SG, Arevalo-Serrano J, Abad VO, Recarte PP, Carreras AG, Pastor GN, et al. Platelet count in first trimester of pregnancy as a predictor of perinatal outcome. Open Access Maced J Med Sci. 2017;5:27-32. https://doi.org/10.3889/oamjms.2017.013
21. Stokes KY, Granger DN. Platelets: a critical link between inflammation and microvascular dysfunction. J Physiol. 2012;590:1023-34.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260604

# Factors associated with unsatisfactory ratings among medical courses in the 2025 Brazilian National Medical Education Assessment Exam: a cross-sectional study

Paola Soledad Mosquera[1]*, Alicia Matijasevich[1], Ivan Wilson Hossni Dias[1], Cristiane de Jesus Almeida[1], Mário César Scheffer[1]

## SUMMARY

**OBJECTIVE:** This study considers institutional and contextual characteristics in an analysis of factors associated with unsatisfactory ratings among medical courses in the 2025 Brazilian National Medical Education Assessment Exam.

**METHODS:** We analyzed secondary data from participating medical courses, with an interest in unsatisfactory ratings (1 or 2 out of 5). To measure associations of unsatisfactory ratings with course characteristics (i.e., institution type, location, existence duration, applicant-vacancy ratio, student-faculty ratio), logistic regression models estimated crude and adjusted odds ratios, with 95%CIs.

**RESULTS:** Of 350 courses (119 public, 231 private), one third (n=107) underperformed. More public (95.8%) than private (55.8%) courses performed satisfactorily. In the multivariate model, courses in the 2nd and 3rd quartiles of the applicant-vacancy ratio had lower odds of receiving an unsatisfactory rating (aOR 0.47; 95%CI 0.23–0.93 and aOR 0.32; 95%CI 0.15–0.67). Courses with less than 10 years of existence had higher odds of an unsatisfactory rating (aOR 2.13; 95%CI 1.01–4.47). Higher odds were observed in the North (aOR 5.41; 95%CI 1.70–17.18), whereas lower odds were found in the South (aOR 0.39; 95%CI 0.15–0.98), compared with the Southeast. Private courses had higher odds of an unsatisfactory rating (aOR 10.93; 95%CI 2.68–44.53).

**CONCLUSION:** Results showed inequalities in medical course ratings by institution type, existence duration, applicant-vacancy ratio, and geographical location, indicating the need to improve regulatory policies and processes for evaluating medical education in Brazil.

**KEYWORDS:** Medical schools. Medical education. Academic performance. Educational measurement. Brazil.

## INTRODUCTION

Over 4,000 medical schools worldwide offer medical education[1], reflecting diverse educational systems and healthcare needs. In Brazil, the number of medical programs increased from 252 to 494 over the past decade[2], consolidating the country's position as having the second most medical schools behind India[1]. However, course heterogeneity regarding admission criteria, infrastructure, and training duration, influenced by regional, political, and economic factors, has sparked global discussions about the capacity and quality of medical education[3,4]. The discourse considers the potential adverse implications for health systems, patient safety, and responses to population health needs.

The evaluation of institutional quality in medical education includes accreditation processes, management-related aspects, pedagogical design, the institutional infrastructure required for medical education[5], and performance measurement systems based on standardized exams applied to students[6,7]. Additionally, faculty evaluations[8], governance structures, educational resources[9], and clinical practice environments[10] are indirect indicators of institutional quality. Unsatisfactory results tend to indicate persistent structural weaknesses in medical courses that may require regulatory actions and educational improvement.

In Brazil, in 2025, the Ministry of Education (MEC) unified the assessment instruments of the National Student Performance Exam (ENADE) and the National Residency

[1]Universidade de São Paulo, Faculdade de Medicina, Departamento de Medicina Preventiva, Programa de Pós-Graduação em Saúde Coletiva – São Paulo (SP), Brazil.

*Corresponding author: paolamosquera@alumni.usp.br

Conflicts of interest: the authors declare there is no conflicts of interest. Funding: this study was supported by the São Paulo Research Foundation (FAPESP; grant No. 2023/10124-0). Alicia Matijasevich (312746/2021-0) and Mário César Scheffer (305433/2025-3) are recipients of research productivity fellowships from the National Council for Scientific and Technological Development (CNPq). Ivan Wilson Hossni Dias (2024/22867-0) and Paola Soledad Mosquera (2024/17428-8) received postdoctoral fellowships from FAPESP. Cristiane de Jesus Almeida (2026/00940-3) is supported by a Technical Training Fellowship (TT-5) from FAPESP.

Received on June 10, 2026. Accepted on July 13, 2026.

Scientific Editor: Roseli Nomura

Exam (ENARE), giving rise to the Brazilian National Medical Education Assessment Exam (ENAMED)[11], which provides a standardized assessment of course-level performance and supports the admission process for medical residency. Although ENAMED does not encompass all dimensions involved in the quality of medical education, it offers a national indicator for evaluating educational outcomes. Graduands and graduates of medical course participants in ENARE are eligible. Among the 350 medical course participants in the inaugural ENAMED, 107 underperformed, prompting the MEC to supervisory actions[12].

Multilevel analysis using data from ENADE and the Higher Education Census (2016–2018) indicated that although individual student characteristics explained some of the variation in performance, course-related attributes significantly influenced the results[13]. Several factors have been identified as potential determinants of the performance of medical courses under ENADE, including institutional characteristics, training capacity-related aspects, human resource availability, and courses' territorial context[14,15]. However, these factors are underexplored in an integrated manner, particularly in analyses that adopt the course as the research unit.

Given the substantial expansion of course offerings, particularly in the private sector, this study recognizes that understanding the factors associated with obtaining unsatisfactory course ratings is relevant to improving regulatory and evaluation strategies. The research aimed to analyze factors associated with unsatisfactory ENAMED 2025 ratings among medical courses, considering their institutional and contextual characteristics.

## METHODS

### Data sources and study variables

This cross-sectional study used secondary data from medical courses participating in the first ENAMED, with the medical course as the unit of analysis. Course information was obtained from the ENAMED 2025 database[11] and the Higher Education Census ("Courses" module)[16], both provided by the National Institute for Educational Studies and Research Anísio Teixeira (INEP), and from the National Register of Higher Education Courses and Institutions (e-MEC)[17]. Population estimates of municipalities housing medical schools were based on data from the Brazilian Institute of Geography and Statistics[18].

The following course characteristics were analyzed: institution type (public vs. private); location (north, northeast, southeast, south, or central-west) and municipal population (>300,000 or <300,000 inhabitants); existence duration (≥21, 11–20, or ≤10 years); applicant-vacancy ratio (i.e., admission competition), estimated as the total number of applicants divided by the total number of vacancies in 2020, considering that ENAMED 2025 participants were admitted to medical courses that year; and student-faculty ratio, estimated based on the total number of enrollees and faculty in 2023, subject to the availability of these data at the time of analysis. This study classified courses offered free of charge as public and those that charged monthly fees as private, regardless of whether the latter were subsidized. This operational classification was adopted because the primary interest was not the legal nature of the institution, but whether access to the course required tuition payment. Municipal population and existence duration were defined based on data or estimates corresponding to 2025, the ENAMED reference year.

The outcome of interest was whether a medical course received an unsatisfactory ENAMED 2025 rating. According to INEP methodology, course ratings are derived from the performance of graduating students with valid examination results, including the proportion achieving a proficient performance level. Courses with fewer than 10 valid participants do not receive a rating. The procedures used to estimate student proficiency and derive course ratings are described in official INEP technical documentation[19,20]. Ratings range from 1 to 5 and were grouped as unsatisfactory (1–2) and satisfactory (3–5)[19].

### Statistical analysis

Course characteristics were described according to the ENAMED rating and type of the providing institution, considering absolute values (n) and relative frequencies (%). We compared course characteristics across ENAMED rating groups using Fisher's exact and chi-square tests for categorical variables. The applicant-vacancy and the student-faculty ratios were asymmetrically distributed between courses and thus described by median and interquartile range (IQR: P25–P75) and compared using the Mann-Whitney test. To assess the association between course characteristics and unsatisfactory ENAMED ratings, logistic regression models estimated crude and adjusted odds ratios (aORs), with respective 95%CIs. In the adjusted model, all explanatory variables were included simultaneously based on their theoretical relevance and previous evidence indicating that institutional, contextual, and organizational characteristics may be associated with the performance of medical courses in educational assessments[14,15]. Given its asymmetric distribution and variability across courses, the applicant-vacancy ratio was categorized into quartiles for association analysis. After adjusting the models, we assessed the presence of multicollinearity among the explanatory variables using the Variance Inflation Factor (VIF). We also conducted a sensitivity analysis in which we estimated the model considering only courses rated 1 or 2 versus those rated 4 or 5 and excluding those rated 3. This verified whether inclusion of this category in

the best performing group influenced the association estimates since a rating of 3 represents an intermediate performance.

For the analyses, we used Stata, version 16.0, and adopted a 5% significance level. This study was approved by the Research Ethics Committee of the Medical School of the University of São Paulo (CAAE No. 71626323.8.0000.0068). Artificial intelligence tools were used exclusively for linguistic revision and had no involvement in the generation of scientific content, data analysis, or the findings.

## RESULTS

Three hundred and fifty courses participated in ENAMED 2025, of which 119 and 231 were public and private, respectively (Table 1). Among them, 107 were rated as unsatisfactory. Most public courses (95.8%) achieved a satisfactory rating, whereas 55.8% of private courses did. Indeed, 102 private courses were rated as unsatisfactory (1 or 2), whereas only five public courses were rated as underperforming.

**Table 1.** Characteristics of medical courses by National Medical Education Assessment Exam 2025 rating and institution type in Brazil (n=350 courses).

| Medical courses | ENAMED 2025 rating | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Total | | | Public courses | | | Private courses | | |
| | 1 and 2 | 3, 4, and 5 | p-value[a] | 1 and 2 | 3, 4, and 5 | p-value[a] | 1 and 2 | 3, 4, and 5 | p-value[a] |
| Courses, n (%) | 107 (30.6) | 243 (69.4) | <0.001 | 5 (4.2) | 114 (95.8) | 0.103 | 102 (44.2) | 129 (55.8) | <0.001 |
| Applicant-vacancy ratio, median (IQR) | 3.6 (1.9–6.8) | 17.0 (5.8–65.9) | | 46.8 (40.5–53.2) | 66.9 (44.7–116.0) | | 3.4 (1.9–6.2) | 7.3 (3.8–13.0) | |
| Applicant-vacancy ratio, n (%) | | | | | | | | | |
| 1st quartile (lowest admission competition) | 52 (59.1) | 36 (40.9) | <0.001 | 1 (11.1) | 8 (88.9) | 0.741 | 51 (64.6) | 28 (35.4) | <0.001 |
| 2nd | 33 (37.9) | 54 (62.1) | | 0 (0.0) | 2 (100.0) | | 33 (38.8) | 52 (61.2) | |
| 3rd | 19 (21.6) | 69 (78.4) | | 1 (4.4) | 22 (95.6) | | 18 (27.7) | 47 (72.3) | |
| 4th quartile (highest admission competition) | 3 (3.5) | 84 (96.5) | | 3 (3.5) | 82 (96.5) | | 0 (0.0) | 2 (100.0) | |
| Student-faculty ratio[b], median (IQR) | 10.0 (7.1–14.6) | 5.6 (4.0–7.9) | <0.001 | 8.7 (7.7–8.9) | 4.5 (3.4–5.8) | <0.001 | 10.3 (6.9–14.7) | 7.1 (5.1–10.8) | <0.001 |
| Existence duration[c], n (%) | | | | | | | | | |
| ≥21 years | 27 (19.2) | 114 (80.8) | <0.001 | 0 (0.0) | 63 (100.0) | 0.014 | 27 (34.6) | 51 (65.4) | 0.001 |
| 11–20 years | 27 (25.2) | 80 (74.8) | | 3 (7.7) | 36 (92.3) | | 24 (35.3) | 44 (64.7) | |
| ≤10 years | 53 (52.0) | 49 (48.0) | | 2 (11.8) | 15 (88.2) | | 51 (60.0) | 34 (40.0) | |
| Region, n (%) | | | | | | | | | |
| North | 14 (46.7) | 16 (53.3) | 0.012 | 2 (14.3) | 12 (85.7) | 0.211 | 12 (75.0) | 4 (25.0) | <0.001 |
| Northeast | 26 (30.2) | 60 (69.8) | | 2 (4.8) | 40 (95.2) | | 24 (54.6) | 20 (45.4) | |
| Southeast | 45 (31.9) | 96 (68.1) | | 0 (0.0) | 30 (100.0) | | 45 (40.5) | 66 (59.5) | |
| South | 8 (13.8) | 50 (86.2) | | 1 (5.3) | 18 (94.7) | | 7 (17.9) | 32 (82.1) | |
| Central-West | 14 (40.0) | 21 (60.0) | | 14 (100.0) | 14 (100.0) | | 14 (66.7) | 7 (33.3) | |
| Location by municipal population size[d], n (%) | | | | | | | | | |
| >300,000 inhabitants | 44 (25.3) | 130 (74.7) | 0.033 | 1 (1.6) | 61 (98.4) | 0.192 | 43 (38.4) | 69 (61.6) | 0.087 |
| <300,000 inhabitants | 63 (35.8) | 113 (64.2) | | 4 (7.0) | 53 (93.0) | | 59 (49.6) | 60 (50.4) | |
| Institution type, n (%) | | | | | | | | | |
| Public | 5 (4.2) | 114 (95.8) | <0.001 | | | | | | |
| Private | 102 (44.2) | 129 (55.8) | | | | | | | |

Source: Authors; National Institute for Educational Studies and Research Anísio Teixeira/Ministry of Education. ENAMED: National Medical Education Assessment Exam; IQR: interquartile range. Note: In 2020, a medical course failed to respond to the Higher Education Census; we obtained its data from the 2021 census. [a]p-values estimated using Mann-Whitney, Fisher's exact, and chi-squared tests, as appropriate. [b]Student-faculty ratio was based on information from the 2023 Higher Education Census. [c]Course existence duration is the period between its inaugural year and 2025. [d]Population estimates for 2025 were obtained from the Brazilian Institute of Geography and Statistics.

The applicant-vacancy ratio was significantly lower among underperforming courses (median: 3.6; IQR: 1.9–6.8) than among those with a satisfactory rating (median: 17.0; IQR: 5.8–65.9). Public courses had a higher applicant-vacancy ratio than private courses. Competition for admission to private courses was significantly lower among those with an unsatisfactory rating (median: 3.4 vs. 7.3). The student-faculty ratio was significantly higher among courses with an unsatisfactory rating (median: 10.0; IQR: 7.1–14.6) than among those with a satisfactory rating (median: 5.6; IQR: 4.0–7.9). Public courses had a lower student-faculty ratio than private courses. In both types of institution, courses with an unsatisfactory rating had more students per faculty member than courses with a satisfactory rating.

The proportion of courses with an unsatisfactory rating increased as existence duration decreased, with 19.2% of courses operating for 21 years or more, 25.2% of those operating for 11–20 years, and 52.0% of those operating for up to 10 years receiving unsatisfactory ratings. Conversely, the proportion of courses with a satisfactory rating was higher among older programs. No public course in existence for 21 years or more (n=63) received an unsatisfactory rating, but 11.8% (n=2) of those less than 10 years old (n=17) did. Among the oldest private courses, 34.6% were rated 1 or 2, and 65.4% were rated 3–5.

The proportion of underperforming courses was 25.3% in schools located in populous municipalities and 35.8% in those in smaller municipalities (p=0.033). No significant differences were observed by institution type after stratification. The south had the lowest proportion of underperforming courses (13.8%), followed by the northeast (30.2%), southeast (31.9%), central-west (40%), and north (46.7%; Table 1).

The multivariate model revealed an association of applicant-vacancy ratio, existence duration, location, and institution

**Table 2.** Adjusted logistic regression analysis of factors associated with an unsatisfactory rating in National Medical Education Assessment Exam in Brazil (n=350 courses).

| Course characteristics | ENAMED rating of 1 or 2 | | | |
|---|---|---|---|---|
| | OR (95%CI) | p-value | aOR (95%CI) | p-value |
| Applicant-vacancy ratio | | | | |
| 1st quartile (lowest admission competition) | 1 | <0.001 | 1 | 0.011 |
| 2nd | 0.42 (0.23–0.77) | | 0.47 (0.23–0.93) | |
| 3rd | 0.19 (0.10–0.36) | | 0.32 (0.15–0.67) | |
| 4th quartile (highest admission competition) | 0.02 (0.01–0.08) | | 0.21 (0.03–1.18) | |
| Student-faculty ratio | 1.09 (1.05–1.13) | <0.001 | 1.02 (0.99–1.06) | 0.138 |
| Existence duration | | | | |
| ≥21 years | 1 | <0.001 | 1 | 0.050 |
| 11–20 years | 1.42 (0.77–2.61) | | 0.90 (0.43–1.86) | |
| ≤10 years | 4.5 (2.57–8.08) | | 2.13 (1.01–4.47) | |
| Region | | | | |
| North | 1.86 (0.83–4.15) | 0.017 | 5.41 (1.70–17.18) | 0.002 |
| Northeast | 0.92 (0.51–1.65) | | 1.64 (0.78–3.42) | |
| Southeast | 1 | | 1 | |
| South | 0.34 (0.14–0.77) | | 0.39 (0.15–0.98) | |
| Central-West | 1.42 (0.66–3.05) | | 1.97 (0.72–5.38) | |
| Location by municipal population size | | | | |
| >300,000 inhabitants | 1 | 0.034 | 1 | 0.305 |
| <300,000 inhabitants | 1.64 (1.03–2.60) | | 1.39 (0.73–2.63) | |
| Institution type | | | | |
| Public | 1 | <0.001 | 1 | 0.001 |
| Private | 18.02 (7.09–45.80) | | 10.93 (2.68–44.53) | |

ENAMED: National Medical Education Assessment Exam; aOR: adjusted odds ratio; CI: confidence interval; OR: odds ratio. Note: Variance inflation factors (VIFs) for the adjusted model indicated no evidence of problematic multicollinearity (mean VIF=1.69; range: 1.16–3.39).

type with courses' ENAMED ratings (Table 2). As the applicant-vacancy ratio increased, the odds of receiving an unsatisfactory rating decreased. Compared to the first quartile comprising institutions with the lowest applicant-vacancy ratio, the second and third quartiles showed 53% (aOR 0.47; 95%CI 0.23–0.93) and 68% (aOR 0.32; 95%CI 0.15–0.67) reductions in the odds of an unsatisfactory rating, respectively. Courses operating for less than 10 years had higher odds of receiving an unsatisfactory rating than those operating for 21 years or more (aOR 2.13; 95%CI 1.01–4.47). Courses offered in the North had higher odds of receiving an unsatisfactory rating (aOR 5.41; 95%CI 1.70–17.18), whereas those in the South had lower odds (aOR 0.39; 95%CI 0.15–0.98) than those in the Southeast. Private courses had higher odds of being rated as unsatisfactory than public courses (aOR 10.93; 95%CI 2.68–44.53). The 95%CI was wide, possibly due to the small number of events among public courses. No evidence of problematic multicollinearity (mean VIF=1.69) was observed among the independent variables included in the adjusted model.

The sensitivity analysis restricted to comparing courses rated 1 or 2 with those rated 4 or 5 yielded results consistent with the main model, with no change in the direction of the associations. Although some 95%CIs were wider because of the reduced sample size, the overall pattern of findings remained unchanged (Table 3).

The results of ENAMED 2025 revealed that 30% of participating courses underperformed. Private courses, those operating for less than 10 years, those with a lower applicant-vacancy ratio, and those offered in the North had higher odds of receiving an unsatisfactory rating.

Prior studies conducted based on ENADE anticipated worse results, on average, for private (vs. public) medical schools[21,22]. The hegemony of private medical schools in Brazil (where 72% of all medical courses in 2024 were private)[2] was accompanied

**Table 3.** Adjusted logistic regression of factors associated with an unsatisfactory rating in National Medical Education Assessment Exam 2025, comparing scores 1–2 versus 4–5, Brazil (n=270).

| Course characteristics | ENAMED rating of 1 or 2 | | | |
|---|---|---|---|---|
| | OR (95%CI) | p-value | aOR (95%CI) | p-value |
| Applicant-vacancy ratio | | | | |
| 1st quartile (lowest admission competition) | 1 | <0.001 | 1 | 0.001 |
| 2nd | 0.42 (0.23–0.77) | | 0.38 (0.14–1.02) | |
| 3rd | 0.19 (0.09–0.36) | | 0.15 (0.05–0.41) | |
| 4th quartile (highest admission competition) | 0.02 (0.01–0.08) | | 0.11 (0.02–0.69) | |
| Student-faculty ratio | 1.09 (1.05–1.13) | <0.001 | 1.03 (0.98–1.09) | 0.181 |
| Existence duration | | | | |
| ≥21 years | 1 | <0.001 | 1 | 0.028 |
| 11–20 years | 1.42 (0.77–2.61) | | 1.13 (0.47–2.72) | |
| ≤10 years | 4.56 (2.57–8.09) | | 3.56 (1.33–9.52) | |
| Region | | | | |
| North | 1.86 (0.83–4.15) | 0.017 | 11.89 (2.11–67.05) | 0.007 |
| Northeast | 0.92 (0.51–1.65) | | 1.75 (0.63–4.86) | |
| Southeast | 1 | | 1 | |
| South | 0.34 (0.14–0.77) | | 0.33 (0.10–1.01) | |
| Central-West | 1.42 (0.66–3.05) | | 1.45 (0.38–5.56) | |
| Location by municipal population size | | | | |
| >300,000 inhabitants | 1 | 0.034 | 1 | 0.288 |
| <300,000 inhabitants | 1.64 (1.03–2.60) | | 1.55 (0.69–3.48) | |
| Institution type | | | | |
| Public | 1 | <0.001 | 1 | <0.001 |
| Private | 18.02 (7.09–45.80) | | 25.27 (5.33–119.67) | |

ENAMED: National Medical Education Assessment Exam; aOR: adjusted odds ratio; CI: confidence interval; OR: odds ratio.

by a significant increase in the offering of physician services and the relative geographical deconcentration of professionals[23]. However, the present findings indicate substantial heterogeneity within the private sector, with a considerable proportion of private courses receiving unsatisfactory ENAMED ratings (44.2%). These results suggest that some private medical courses may require closer regulatory monitoring and targeted quality-improvement strategies.

In line with our results, previous studies have demonstrated better performance of courses in capital cities[13], as well as those in the south, and worse performance of courses in the north[14]. International evidence has shown that medical schools and other factors may contribute to the establishment of physicians in underserved regions and/or those that are remote from big cities[23]. However, regional differences may reflect broader inequalities in educational and healthcare infrastructure, availability of clinical training environments, and faculty recruitment and retention. The regionalization of medical education remains an important strategy for expanding access to training and addressing workforce shortages, but it should be accompanied by adequate investments and quality-monitoring mechanisms. Further studies are needed to assess whether limitations in infrastructure, clinical training environments, and faculty availability are related to poor ENAMED performance.

This study found an association of greater admission competition (applicant-vacancy ratio) with a lower chance of an unsatisfactory rating in ENAMED. Medicine is among the undergraduate courses with the highest admission competition in Brazil, but the applicant-vacancy ratio varies widely among institutions. The literature identifies the diversity of admission processes and recruitment formats based on standardized tests, previous school performance, affirmative policies, and other factors as a methodological challenge to analyzing the relationship between admission and academic performance[24]. In this context, discussing the improvement of admission mechanisms in medical schools is necessary, with consideration to both performance in evaluations and the training of professionals to foster their development of capabilities appropriate to medical practice.

Relatively new medical courses performed poorly in ENAMED. Many were authorized in the last decade through expansion policies such as the Lei Mais Médicos (More Physicians Act 2013)[25] and, in some cases, by court decisions that required the MEC to process opening requests. Shorter existence duration may reflect lower institutional maturity, including less consolidated educational projects, faculty structures, clinical training environments, and integration with health services. These findings reinforce the importance of not only meeting initial requirements for course authorization but also ensuring adequate conditions for sustaining educational quality over time.

Overall, the findings were robust across sensitivity analyses. The exclusion of courses with concept 3 yielded consistent estimates, with no changes in the direction of associations, although with wider confidence intervals.

This study has limitations. Its cross-sectional design does not allow causal inferences. The analysis relied on secondary databases and used the medical course as the unit of analysis, precluding individual-level interpretations. In addition, relevant institutional characteristics, such as clinical training environments, preceptorship, and teaching-service integration, were unavailable. Furthermore, the classification of institutions as public or private may not fully capture the heterogeneity of the private sector, potentially introducing residual confounding. Finally, estimates related to institution type should be interpreted cautiously due to the small number of unsatisfactory ratings among public courses, which may have reduced the precision and stability of the corresponding estimates.

## CONCLUSION

Approximately one third of ENAMED 2025 courses received an unsatisfactory rating. Private institutions, shorter course existence duration, less competitive admission processes (lower applicant-vacancy ratios), and location in the North region were associated with this outcome. These findings may help strengthen regulatory policies in Brazil and identify courses that warrant closer monitoring and targeted quality-improvement strategies, although ENAMED results should be interpreted alongside other indicators of medical education quality.

## AUTHORS' CONTRIBUTIONS

**PSM:** Data curation, Formal Analysis, Writing – original draft. **AM:** Formal Analysis, Supervision, Writing – review & editing. **IWHD:** Data curation, Formal Analysis, Writing – review & editing. **CJA:** Data curation, Formal Analysis, Writing – review & editing. **MCS:** Conceptualization, Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

The original datasets used in this study are publicly available from their respective sources. The merged analytical dataset, including derived variables created for the analyses, is available from the corresponding author upon reasonable request.

## REFERENCES


1. World Federation for Medical Education, FAIMER. About the world directory of medical schools 2024. 2024. [cited on 2026 Mar 3]. Available from: https://www.wdoms.org/about/
2. Scheffer M. Demografia Médica no Brasil 2025. Brasília (DF): Ministério da Saúde; 2025.
3. Jeyaraju M, Linford H, Bosco Mendes T, Caufield-Noll C, Tackett S. Factors leading to successful performance on U.S. national licensure exams for medical students: a scoping review. Acad Med. 2023;98(1):136-48.
4. Ferreira T, Collins AM. Too many, too soon? Challenges in medical school expansion in the United Kingdom. Hum Resour Health. 2025;23(1):72.
5. World Federation for Medical Education. WFME basic medical education standards 2020. 2020. [cited on 2026 Mar 3]. Available from: https://wfme.org/wp-content/uploads/2022/03/WFME-BME-Standards-2020.pdf
6. United States Medical Licensing Examination. USMLE updates and research. [cited on 2026 Mar 3]. Available from: https://www.usmle.org/
7. General Medical Council. Medical licensing assessment. [cited on 2026 Mar 3]. Available from: https://www.gmc-uk.org/education/medical-licensing-assessment
8. Shrivastava SR, Shrivastava PS. Strengthening the process of evaluation of teaching quality in medical education. Apollo Med. 2021;18(1_suppl):S17-20.
9. Arja SB, White BA, Fayyaz J, Thompson A. The impact of accreditation on continuous quality improvement process in undergraduate medical education programs: a scoping review. MedEdPublish. 2024;14:13.
10. Sellberg M, Palmgren PJ, Möller R. A cross-sectional study of clinical learning environments across four undergraduate programs using the undergraduate clinical education environment measure. BMC Med Educ. 2021;21(1):258.
11. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Brasil). Exame Nacional de Avaliação da Formação Médica (Enamed) 2025. 2026. [cited on 2026 Mar 3]. Available from: https://www.gov.br/inep/pt-br/areas-de-atuacao/avaliacao-e-exames-educacionais/enamed
12. Ministério da Educação (Brasil). Enamed: divulgadas avaliação dos cursos de medicina e medidas de supervisão 2026. 2026. [cited on 2026 Mar 3]. Available from: https://www.gov.br/mec/pt-br/assuntos/noticias/2026/janeiro/enamed-divulgadas-avaliacao-dos-cursos-de-medicina-e-medidas-de-supervisao
13. Santos PS, Oliveira SV, Becker KL. Determinants of academic performance in Brazilian public higher education: a multilevel analysis. Planej Polít Públicas. 2025;68:189-212.
14. Fatel QH, Silva FG, Avena KM, Menezes RC, Andrade BB, Quintanilha LF. Shaping tomorrow's doctors: the impact of socioeconomic and institutional factors on medical education quality in Brazil: an ecological study. J Med Educ Curric Dev. 2025;12:23821205251360521.
15. Neto TA, Silva Fucuta Pereira P, Nogueira ML, Pereira Gody JM, Moscardini AC. Factors that affect the national student performance examination grades of Brazilian undergraduate medical programs. GMS J Med Educ. 2018;35(1):Doc8.
16. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (Brasil). Censo da Educação Superior – Módulo Cursos 2020. 2020. [cited on 2026 Mar 3]. Available from: https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-da-educacao-superior/resultados
17. Ministério da Educação (Brasil). Cadastro Nacional de Cursos e Instituições de Educação Superior: Cadastro e-MEC 2025. [cited on 2026 Jan 20]. Available from: https://emec.mec.gov.br/emec/nova
18. Instituto Brasileiro de Geografia e Estatística. Estimativas da população para 2025. Rio de Janeiro (RJ): IBGE; 2026. [cited on 2026 Mar 3]. Available from: https://www.ibge.gov.br/estatisticas/sociais/populacao/9103-estimativas-de-populacao.html
19. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Nota Técnica nº 19/2025/CGAFM/DAES-INEP: apresentação dos procedimentos adotados para estimativa do desempenho esperado pelos participantes do Enamed, edição 2025. Brasília, DF: INEP; 2025.
20. Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira. Nota Técnica nº 40/2025/CEI/CGGI/DAES-INEP. Brasília (DF): INEP; 2025.
21. Scheffer MC, Poz MR. The privatization of medical education in Brazil: trends and challenges. Hum Resour Health. 2015;13(1):96.
22. Andrade BB, Villalva-Serra K, Menezes RC, Quintanilha LF, Avena KM. For-profit growth and academic decline: a retrospective nationwide assessment of Brazilian medical schools. Front Med (Lausanne). 2025;12:1617885.
23. Scheffer M, Mosquera P, Cassenote A, McPake B, Russo G. Brazil's experiment to expand its medical workforce through private and public schools: impacts and consequences of the balance of regulatory and market forces in resource-scarce settings. Glob Health. 2025;21(1):14.
24. Patterson F, Knight A, Dowell J, Nicholson S, Cousans F, Cleland J. How effective are selection methods in medical education? A systematic review. Med Educ. 2016;50(1):36-60.
25. Brasil. Lei nº 12.871, de 22 de outubro de 2013. Institui o Programa Mais Médicos e dá outras providências. Diário Oficial da União; 2013.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260646

# Osteopathic manual therapy in fibromyalgia: a pilot study

Aziz Dengiz[1], Elif Gur Kabul[2], Gulsum Tikac[3]*, Bilge Basakci Calik[4], Uğur Karasu[5]

## SUMMARY

**OBJECTIVE:** The aim of this pilot study was to evaluate the preliminary efficacy and feasibility of integrating Osteopathic Manual Therapy Techniques combined with Clinical Pilates on pain, symptom severity, and sleep quality in patients with fibromyalgia.

**METHODS:** Twenty women with fibromyalgia were randomized into Osteopathic Manual Therapy Techniques+Clinical Pilates (n=11) and Clinical Pilates (n=9) groups. The protocol lasted 6 weeks (3 sessions/week). Osteopathic Manual Therapy Techniques included rib raising, suboccipital myofascial release, Sternocleidomastoid (SCM) mobilization, diaphragm/mediastinum techniques, and visceral mobilizations. Primary outcomes were the Widespread Pain Index, Symptom Severity Scale, and Pressure Pain Threshold. Sleep was assessed via the Pittsburgh Sleep Quality Index.

**RESULTS:** Both groups showed significant intra-group improvements ($p<0.05$). Within-group analyses revealed that the Osteopathic Manual Therapy Techniques+Clinical Pilates group achieved significant improvements in Widespread Pain Index ($p=0.003$), Symptom Severity Scale ($p=0.005$), and Pittsburgh Sleep Quality Index ($p=0.007$). However, in the between-group analysis of change scores, the Osteopathic Manual Therapy Techniques+Clinical Pilates group demonstrated significantly superior improvements exclusively in sleep quality (Pittsburgh Sleep Quality Index, $p=0.034$) and maximal Pressure Pain Threshold at the trapezius ($p=0.035$) compared to the Clinical Pilates-only group, while between-group differences for Widespread Pain Index and Symptom Severity Scale did not reach statistical significance ($p>0.05$).

**CONCLUSION:** Integrating Osteopathic Manual Therapy Techniques into exercise-based rehabilitation presents promising preliminary evidence for enhancing central sensitization symptoms and sleep quality in fibromyalgia patients, though larger trials are needed.

**KEYWORDS:** Fibromyalgia. Manipulative medicine. Osteopathic medicine. Physical therapy modalities. Exercise therapy.

## INTRODUCTION

Fibromyalgia (FM) is a chronic syndrome characterized by widespread musculoskeletal pain accompanied by fatigue, sleep, memory, and mood issues[1]. FM is fundamentally characterized by central sensitization and a profound imbalance of the autonomic nervous system, where persistent sympathetic overactivity and suppressed parasympathetic function lead to abnormal pain processing. This autonomic dysregulation results in systemic vasoconstriction and subsequent ischemic events within the tissues, which are hypothesized to be the primary drivers of widespread musculoskeletal pain, fatigue, and hypersensitivity[2]. Osteopathic Manual Therapy (OMT) is hypothesized to address these mechanisms by employing specific manual techniques aimed at modulating autonomic tone—potentially downregulating the sympathetic "fight or flight" response and enhancing parasympathetic activity—while simultaneously improving local blood flow to reverse tissue ischemia and soothe the hypersensitive central nervous system[3]. The present study is important in terms of applying techniques that aim to affect the autonomic nervous system and measuring the effectiveness of these techniques.

While Clinical Pilates (CP) is effective, pain-related barriers often limit results. Osteopathic Manual Therapy Techniques (OMTT) aim to balance the autonomic nervous system[2]. We aimed to investigate the preliminary therapeutic benefits and clinical applicability of osteopathic manual techniques for pain and symptom management in FM patients, thereby providing feasibility data for larger, multi-center trials.

## METHODS

This study was approved by the Local Ethics Committee (Decision no: 09, Date: 05.30.2023) and performed in accordance with

[1]Gaziantep Islam Science and Technology University, Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation – Gaziantep, Türkiye.
[2]Usak University, Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation – Uşak, Türkiye.
[3]Siirt University, Faculty of Health Sciences, Department of Physiotherapy and Rehabilitation – Siirt, Türkiye.
[4]Pamukkale University, Faculty of Physiotherapy and Rehabilitation, Department of Orthopedic Rehabilitation – Denizli, Türkiye.
[5]Pamukkale University, Faculty of Medicine, Department of Internal Medicine – Denizli, Türkiye.
*Corresponding author: fztgulsumtikac@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on April 12, 2026. Accepted on July 19, 2026.
Scientific Editor: José Maria Soares Júnior

the Declaration of Helsinki. All participants provided written informed consent prior to inclusion.

In this randomized controlled clinical trial, conducted in accordance with the Consolidated Standards of Reporting Trials (CONSORT) guidelines, participants were assigned to either the Intervention Group (Osteopathic Manual Therapy Techniques+Clinical Pilates, OMTT+CP) or the Control Group (CP only) via a simple, unconstrained computer-generated randomization sequence (SPSS v.22.0) with a 1:1 allocation ratio.

The study included 28 women with FM. FM was diagnosed by the same rheumatologist. Following eight dropouts due to personal reasons in the OMTT+CP group (n=3) and hectic work schedules in the CP group (n=5), the remaining 20 women were included in the final analysis, as detailed in the CONSORT flow diagram (Figure 1). Inclusion criteria: (1) Female, aged 18–65 years; (2) Diagnosed with FM based on the 2016 American College of Rheumatology (ACR) criteria; (3) Stable medication dose for at least 6 months. Exclusion criteria: (1) Presence of neurological, orthopedic, cardiopulmonary, or inflammatory diseases; (2) Cognitive impairment; (3) Pregnancy; (4) Participation in regular exercise or rehabilitation programs for the 3 months in the preceding year.

## Interventions

Both groups performed CP sessions three times weekly for 6 weeks.

Clinical Pilates (CP): CP sessions were conducted as group therapy under the supervision of a single physiotherapist. Prior to the intervention, participants were instructed in the five key elements of Pilates: lateral costal breathing, centering (focus), and the neutral alignment of the rib cage, shoulder girdle, and head/neck. Each 60-min session comprised a 10-min warm-up, 40 min of core-stabilization exercises, and a 10-min cool-down. Exercises initially consisted of 6–8 repetitions, with progression implemented at weeks 2 and 4 by increasing repetitions or adding resistance via elastic bands (e.g., advancing from yellow to red or green Theraband) based on individual tolerance.

Osteopathic Manual Therapy (OMTT): The intervention group received 30-min OMTT sessions prior to CP. All OMTT interventions were administered as a standardized protocol in every single session (three times weekly for 6 weeks, totaling 18 sessions). To eliminate inter-examiner variability, all sessions were delivered by a single, certified osteopathic practitioner/physiotherapist with over 5 years of clinical experience in manipulative medicine. The techniques were applied in a strict, uniform chronological sequence in accordance with a standardized protocol adapted from Nicholas and Nicholas. The specific clinical sequence and precise application durations were executed as follows:

1. Rib raising technique: Applied for approximately 2 min with the patient in the supine position while performing bilateral oscillation over the ribs.
2. Suboccipital myofascial technique: Applied for an average of 4–5 min in the supine position, utilizing finger semi-flexion between the occiput and axis until suboccipital relaxation was achieved, followed by light traction.
3. Bilateral SCM mobilization: Executed for an average of 2–3 min with anterior-posterior mobilization of the right and left SCM muscles in the supine position.
4. Mediastinum technique: Performed for an average of 8–10 repetitions with the patient in the side-lying position, applying synchronized multi-directional pressures to the sternum and thoracic region.
5. Diaphragm mobilization techniques: Administered via a three-step sequence (transverse costal mobilization for 1–2 min in supine with knees bent, respiratory-resisted

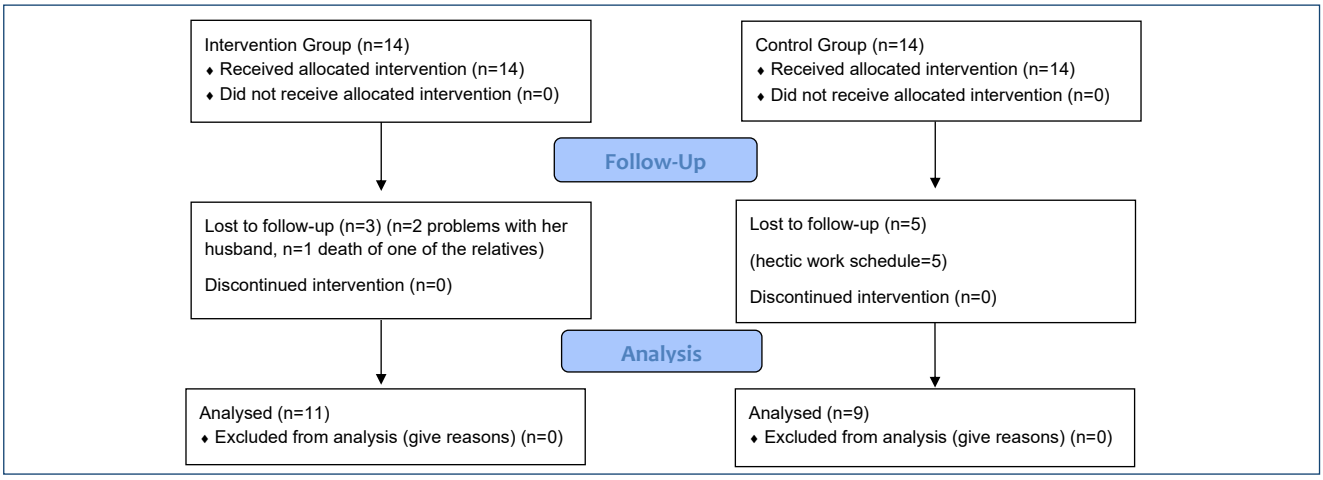



**Figure 1.** Consolidated Standards of Reporting Trials flow diagram of the study.

manual tracking under the costal arch in supine, and manual costal border resistance in a sitting position).

6. Sacrum mobilization technique: Applied for 2 min with the patient in the prone position, delivering anterior-cranial oscillations at a rate of 150–180 beats per minute over the lower third of the sacrum.
7. Celiac and mesenteric ganglion oscillation technique: Performed for an average of 4 min in the supine position with knees flexed, applying manual oscillation therapy under the xiphoid process (celiac ganglion) and 2 cm caudally (mesenteric ganglion)[4].

### Assessments

Evaluations were performed at baseline and after 6 weeks by a blinded researcher.

Primary outcomes: Pain intensity and symptom severity were measured using the Widespread Pain Index (WPI), Symptom Severity Scale (SSS), and Pressure Pain Threshold (PPT).

Secondary outcomes: Functional status (FIQ), sleep quality (PSQI), fatigue (MAF), and quality of life (SF-36) were assessed using validated Turkish versions of the respective scales. Exercise beliefs were evaluated with the Exercise Benefits/Barriers Scale (Exercise Beliefs Questionnaire-EBBS). Anxiety (BAI) and biopsychosocial status (Biopsychosocial Questionnaire [BETY-BQ]) were also recorded.

The WPI records widespread pain by questioning 19 specific regions (0–19 range). Symptom severity was measured by the SSS (0–12), which rates fatigue, sleep quality, and cognitive symptoms (0–3 points), and identifies the presence of headache, abdominal pain, and depression[5]. PPT were assessed at four bilateral sites (trapezius, lateral epicondyle, quadriceps femoris, and medial knee) using a digital algometer (JTECH Medical™ Commander Echo). The 1-$cm^2$ rubber-tipped probe was applied perpendicularly to the skin, with pressure increased steadily until the participant reported the pain threshold. For each site, both maximal and mean force values ($kg/cm^2$) were recorded[6].

The FIQ, developed by Salaffi et al. to assess functional status in FM patients, comprises 10 items with a total score up to 100, where higher results indicate more severe impairment[7]. Sleep quality was evaluated using the PSQI, which scores seven components to provide a total range of 0–21, with higher values reflecting poorer sleep quality[8]. Fatigue was quantified via the 16-item MAF scale, measuring its severity and impact on daily activities[9]. Health-related quality of life was measured using the SF-36, where higher physical and mental component scores represent better overall well-being[10]. Additionally, the 43-item EBBS was used to determine exercise perceptions and barriers[11], while the BAI assessed anxiety severity, with higher scores indicating increased anxiety levels[12]. Finally, the BETY-BQ evaluated the biopsychosocial status across 30 items, where higher scores signify lower health status[13].

### Statistical analysis

Sample size calculation was performed via G*Power 3.1. To achieve 80% power with a 5% type-1 error rate ($\alpha$=0.05) and an effect size of d=1.09[14], a minimum of 24 subjects (12 per group) was required. To account for a potential drop-out rate during the 6-week follow-up, we intentionally over-recruited and randomized 28 participants (14 per group), exceeding the minimum requirement of 24 subjects. However, due to an unexpected attrition of eight participants during the follow-up period (reasons detailed in Figure 1), the final analysis was completed with 20 participants (11 intervention, 9 control).

Data were analyzed using International Business Machines Statistical Package for the Social Sciences Statistics v.22.0. Normality was checked via the Shapiro-Wilk test. Due to the small sample size (n=20) and the non-normal distribution of several clinical parameters, non-parametric statistical methods were applied. Non-normally distributed continuous variables are expressed as median (min–max). The Wilcoxon signed-rank and Mann-Whitney U tests were used for intra-group and inter-group comparisons, respectively. Between-group comparisons of therapeutic efficacy were performed using change scores (Delta=post - pre) via the Mann-Whitney U test. Although analysis of covariance is widely preferred to control for baseline variances in larger trials, it was not methodologically feasible in this study due to the violation of parametric assumptions (normality and sample size constraints) required for general linear models. Statistical significance was set at $p<0.05$.

## RESULTS

Twenty female patients with FM completed the study (Intervention: n=11; Control: n=9). The participant flow is detailed in the CONSORT diagram (Figure 1). Baseline demographic and disease-related characteristics showed no significant differences between the groups (Mann-Whitney U test, all $p>0.05$). Specifically, the median (min/max) age was 52 (38/62) years for the intervention group and 51.5 (30/63) years for the control group ($p=0.469$). Median height and weight were 1.56 m (1.52/1.66) and 74 kg (56/90) in the intervention group, compared to 1.60 m (1.55/1.67) and 74.5 kg (53/99) in the control group ($p=0.731$ and $p=0.820$, respectively). Furthermore, the median disease duration was 3.5 (1/11) years for the intervention group and 2 (0.5/22) years for the control

group (p=0.360). Although slight baseline discrepancies existed in clinical parameters (such as disease duration and FIQ scores), these differences did not reach statistical significance, and subsequent inter-group evaluations were strictly performed using "change scores" (post-treatment minus baseline) to eliminate any potential confounding effects of minor baseline variances.

Following the treatment, intra-group analyses revealed that the intervention group (OMTT+CP) demonstrated significant improvements in WPI, SSS, overall sleep quality (PSQI), fatigue (MAF), and both the Physical and Mental components of the SF-36. However, no significant changes were observed in FIQ, BETY-BQ, or BAI scores within this group. Conversely, the control group showed significant post-treatment improvements only in MAF and BETY-BQ scores, with no significant changes in PPT values, WPI, SSS, PSQI, or SF-36 components (Table 1).

In inter-group comparisons of the change scores, the intervention group exhibited significantly greater reductions in

**Table 1.** Intragroup results of women with fibromyalgia.

| Variables | Intervention group (n=11) | | | | Control group (n=9) | | | |
|---|---|---|---|---|---|---|---|---|
| | Pre/median (min–max) | Post/median (min–max) | z | p | Pre/median (min–max) | Post/median (min–max) | z | p |
| Maximal PPT | | | | | | | | |
| m. trapezius | 4.60 (3.50/8.10) | 7.70 (4.80/12.70) | | **0.003** | 5.22 (2.90/7.10) | 4.20 (0.80/8.10) | | 0.780 |
| Lateral epicondyle | 4.15 (2.50/10.80) | 6.85 (2.80/10.15) | | **0.010** | 3.15 (1.70/5.80) | 3.55 (0.95/7.45) | | 0.470 |
| m. quadriceps femoris | 8.10 (4.25/13.75) | 11.30 (4.65/19.40) | | **0.041** | 6.22 (4.10/10) | 5.52 (0.9/14.05) | | 0.432 |
| Medial knee | 5.40 (3.20/8.70) | 7.20 (2.60/14.85) | | **0.013** | 3.07 (1.10/9.45) | 2.45 (0.55/9.45) | | 0.055 |
| Mean PPT | | | | | | | | |
| m. trapezius | 4.15 (3.15/7.90) | 7.05 (3.80/11.05) | | **0.004** | 4.72 (2.70/6.50) | 3.80 (0.70/7.25) | | 0.507 |
| lateral epicondyle | 3.90 (2.15/9.05) | 6.45 (2.45/9.35) | | **0.017** | 2.77 (1.45/5.45) | 3.12 (0.75/6.05) | | 0.730 |
| m. quadriceps femoris | 7.4 (4.05/11.85) | 10.80 (4.45/18.50) | | **0.033** | 5.52 (3.60/8.70) | 5.10 (0.75/12.15) | | 0.510 |
| Medial knee | 5 (3.05/8.50) | 6.90 (2.35/14.20) | | 0.068 | 2.50 (0.90/8.75) | 2 (0.55/9.25) | | 0.124 |
| WPI | 16 (9–19) | 11 (0–15) | -2.938 | **0.003*** | 12 (5–19) | 11 (2–19) | -1.069 | 0.285 |
| SSS | 8 (6–11) | 4 (0–10) | -2.807 | **0.005*** | 6.5 (4–12) | 5.5 (1–10) | -1.511 | 0.131 |
| FIQ | 64.75 (43.52–89.10) | 58.03 (31.61–79.80) | -0.800 | 0.424 | 54.31 (18.66–78.77) | 41.12 (5.45–86.24) | -1.572 | 0.116 |
| PSQI (total) | 9 (4–15) | 6 (2–11) | -2.712 | **0.007*** | 7 (2–13) | 7 (3–12) | -0.816 | 0.414 |
| MAF | 36.24 (20.54–44.72) | 26.47 (12.50–40.50) | -2.756 | **0.006*** | 26.16 (12.27–47.72) | 15.85 (9.77–33.68) | -1.992 | **0.046*** |
| SF-36 physical | 139.55 (62.50–237.50) | 180.45 (82.50–320.00) | -2.194 | **0.028*** | 201.07 (40.0–327.5) | 285.36 (200.00–355) | -1.761 | 0.078 |
| SF-36 mental | 160.99 (97.00–259.80) | 215.34 (115.50–306.70) | -2.293 | **0.022*** | 229.54 (77.5–335.5) | 278.37 (193.00–358.00) | -0.135 | 0.893 |
| BAI | 24 (7–43) | 23 (7–46) | -0.119 | 0.906 | 18 (3–31) | 12 (6–19) | 0.000 | 1.000 |
| BETTY total | 56 (28–76) | 55 (32–73) | -0.059 | 0.953 | 43 (21–77) | 22 (10–35) | -2.371 | **0.018*** |
| EBBS—benefits | 70 (61–78) | 69 (53–78) | -0.280 | 0.779 | 72 (61–78) | 72 (61–78) | -0.447 | 0.655 |
| EBBS—barriers | 37 (17–54) | 37 (26–48) | -0.059 | 0.953 | 30 (12–49) | 22 (14–32) | -1.577 | 0.115 |

PPT: Pressure Pain Threshold, m: musculus, WPI: Widespread Pain Index, SSS: Symptom Severity Scale, FIQ: Fibromyalgia Impact Questionnaire, PSQI: Pittsburgh Sleep Quality Index, MAF: Multidimensional Assessment of Fatigue, BAI: Beck Anxiety Inventory, FM: Fibromyalgia, BETTY: Biopsychosocial Questionnaire, EBBS: Exercise Benefits/Barriers Scale, *p<0.05. Wilcoxon Paired-Sample Test. The bold values indicates statistical significance (p<0.05).

sleep quality (PSQI, p=0.034) compared to the control group. Regarding pressure pain thresholds (PPTs), a statistically significant between-group difference was observed only for the maximal PPT at the trapezius muscle (p=0.035). In contrast, the between-group differences for WPI (p=0.389) and SSS (p=0.210) change scores did not reach statistical significance (p>0.05), despite the significant intra-group improvements observed within the intervention cohort. Furthermore, the mean PPT change score for the trapezius demonstrated only a trend toward statistical significance (p=0.088), and no significant differences were found at the other anatomical sites (Table 2).

## DISCUSSION

The primary finding of this pilot study is that while CP alone improves FM symptoms, a multi-modal approach combining CP with OMTT shows promising trends and preliminary superiority in reducing pain and improving sleep quality compared to CP alone. Specifically, the intervention group demonstrated significant improvements in the number of painful areas, PPTs, and sleep quality compared to the control group. OMTT significantly boosts the efficacy of structured exercise.

Current international guidelines consistently identify exercise as the first-line non-pharmacological intervention for FM[15]. Among various exercise modalities, CP is highly recommended for its safety and efficacy in reducing pain, tenderness, and functional limitations. A recent systematic review and meta-analysis by Schaffer et al. confirmed that Pilates significantly improves pain and disability in FM patients[16]. In our study, the significant improvement in symptoms within the control group aligns with these recent findings, confirming that CP remains a robust therapeutic tool. However, the additional gains in the intervention group highlight that exercise alone may not fully address the complex autonomic and neuroendocrine dysfunctions inherent in FM.

**Table 2.** Between-group comparison of change scores (Δ=post−pre) in women with fibromyalgia.

| Variables | Intervention group Δ Median (min–max) | Control group Δ Median (min–max) | Z | p |
|---|---|---|---|---|
| PPT max | | | | |
| m. trapezius | 2.20 (0.20–8.25) | 0.10 (-4.25 to 1.35) | -2.112 | **0.035*** |
| lat. epicondyle | 2.35 (-1.70 to 6.05) | 0.37 (-2.35 to 3.35) | -1.106 | 0.269 |
| m. QF | 2.75 (-1.65 to 11.30) | -0.02 (-5.35 to 4.05) | -0.402 | 0.687 |
| medial knee | 1.25 (-0.60 to 9.45) | -0.55 (-3.75 to 0.85) | -1.358 | 0.175 |
| PPT mean | | | | |
| m. trapezius | 1.55 (-0.55 to 6.90) | -0.07 (-3.80 to 1.85) | -1.709 | 0.088 |
| lat. epicondyle | 2.15 (-1.50 to 6.35) | 0.35 (-2.35 to 2.55) | -1.206 | 0.228 |
| m. QF | 2.90 (-1.55 to 11.00) | 0.07 (-5.15 to 3.45) | -0.654 | 0.513 |
| medial knee | 0.45 (-1.60 to 9.20) | -0.40 (-3.40 to 0.65) | -0.805 | 0.421 |
| WPI | -4 (-13 to -1) | 0 (-11 to 7) | -0.862 | 0.389 |
| SSS | -3 (-9 to 0) | -2 (-4 to 3) | -1.253 | 0.21 |
| FIQ | 8 (-18.78 to 44.68) | 10 (-4.20 to 41.09) | -0.804 | 0.421 |
| BETTY total | 0 (-20 to 15) | 12 (1–20) | -1.617 | 0.106 |
| EBBS benefits | 1 (-10 to 12) | 0 (-4 to 2) | -0.240 | 0.810 |
| EBBS barriers | 0 (-15 to 17) | 5 (-2 to 19) | -0.770 | 0.441 |
| PSQI | 3 (-1 to 8) | 0 (-1 to 2) | -2.123 | **0.034*** |
| MAF | 10 (-1.40 to 18.41) | 12 (-1.91 to 32.50) | 0.000 | 1.000 |
| SF-36 physical | -41 (-110 to 15) | -60 (-117.5 to 10) | -0.738 | 0.461 |
| SF-36 mental | -54 (-123 to 33) | -23 (-148.67 to 77) | -0.794 | 0.427 |
| BAI | 0 (-10 to 23) | 0 (-7.07 to 7.07) | -0.171 | 0.864 |

m. QF: Musculus Quadriceps Femoris; lat. epicondyle: lateral epicondyle; PPT: Pressure Pain Threshold; WPI: Widespread Pain Index; SSS: Symptom Severity Scale; FIQ: Fibromyalgia Impact Questionnaire; MAF: Multidimensional Assessment of Fatigue; BAI: Beck Anxiety Inventory; BETTY: Biopsychosocial Questionnaire; EBBS: Exercise Benefits/Barriers Scale; PSQI: Pittsburgh Sleep Quality Index. Data are presented as median (minimum–maximum) change scores (Δ=post−pre), Between-group comparisons were performed using the Mann-Whitney U test, *p<0.05. The bold values indicates statistical significance (p<0.05).

FM is fundamentally characterized by central sensitization and a disrupted autonomic nervous system (ANS)[17]. Patients typically exhibit a state of "sympathovagal imbalance," marked by sympathetic hyperactivity and reduced parasympathetic tone[18]. OMTT is thought to target these specific pathways to help restore homeostasis. In our protocol, rib raising techniques (T1–T8) and ganglion mobilizations were utilized with the therapeutic intent of downregulating sympathetic tone. Theoretically, increased sympathetic activity leads to regional vasoconstriction and muscle hypoperfusion, which is a primary driver of localized tender points[19]. By potentially enhancing local circulation and oxygenation through these techniques might have contributed to the significant increase in PPT levels observed in our intervention group.

Simultaneously, we focused on facilitating parasympathetic activation via suboccipital myofascial release and sacrum mobilization. A recent meta-analysis by Cerritelli et al. demonstrated that OMTT induces a predominant parasympathetic shift, effectively modulating the ANS and increasing vagal tone[20]. Literature suggests that manual stimulation of fascial tissues may activate free nerve endings that communicate with the insular cortex and hypothalamus, potentially triggering a systemic parasympathetic response[21]. This "bottom-up" modulation is essential for patients with FM, who often show altered central pain processing.

It is worth noting that while the between-group analysis revealed a significant improvement in the maximal PPT of the trapezius muscle, the mean PPT for the same muscle reached only a trend toward significance. This divergence can be methodologically attributed to the localized nature of our pilot sample size and the inherent symptomatic volatility of FM patients. Maximal PPT may respond more rapidly to acute localized therapeutic interventions like suboccipital release and rib raising. Conversely, mean PPT scores are more susceptible to intra-session variability and baseline pain fluctuations, potentially diminishing statistical power in small cohorts.

Our results are strongly supported by recent high-quality trials. Arienti et al. conducted a randomized controlled trial demonstrating that OMTT significantly improves pain intensity and functional status in women with FM[22]. Similarly, Dal Farra et al. concluded that OMT sessions predominantly based on myofascial release play a therapeutic role in rapidly improving functional status and quality of life[23].

Our findings contrast with older studies, such as Coste et al., who found no significant benefit of OMT over sham treatment[24]. We attribute this discrepancy to our treatment frequency and multi-modal design. While that study utilized OMT once weekly as a standalone, joint-focused therapy, our study integrated OMT with CP three times weekly, emphasizing soft tissue mobility. This supports the growing modern consensus that OMT is most effective as an adjunct therapy to exercise rather than a mono-therapy[25].

The marked improvement in sleep quality (PSQI) in the OMTT group is one of our most vital findings. FM-related sleep disturbances are often resistant to standard exercise alone. By potentially encouraging a shift toward parasympathetic tone, OMTT facilitates a state of physiological relaxation, thereby improving sleep onset and overall sleep quality[20].

An intriguing finding in our study was that despite significant improvements in pain distribution (WPI) and symptom severity (SSS) within the intervention group, these changes did not translate into a statistically significant reduction in overall FIQ scores. This apparent discrepancy can be attributed to several factors. First, the FIQ is a comprehensive instrument that evaluates long-term functional status, activities of daily living, and work difficulties over the past week. While OMTT successfully downregulates acute sympathetic hyperreactivity and rapidly modulates localized pain thresholds (PPT), altering a patient's deeply ingrained chronic functional limitations and perceived work difficulty likely requires a longer intervention period than 6 weeks. Second, as a pilot study with a small sample size, the final analysis was likely underpowered to detect smaller clinical changes in multi-dimensional scales like the FIQ, potentially increasing the risk of a Type II error. Nevertheless, the visual trend toward improvement in FIQ scores within the OMTT+CP group provides a promising foundation for future long-term trials.

To our knowledge, this is one of the first studies to specifically investigate the synergistic effects of OMTT and CP in FM within the context of recent neurobiological understandings. Furthermore, the pattern of statistical significance observed between intra-group and inter-group analyses warrants cautious interpretation. While the control group demonstrated significant intra-group improvements in fatigue (MAF) and biopsychosocial status (BETY-BQ)—confirming the well-established benefits of CP on general well-being—the between-group analysis revealed that the synergistic addition of OMTT specifically drove superior outcomes in localized mechanical pain thresholds (maximal PPT at the trapezius) and sleep quality (PSQI). This highlights that while exercise lays a foundational recovery for fatigue and biopsychosocial status, OMTT acts as a powerful targeted accelerator for central sensitization and autonomic sleep-wake regulation. The lack of significant between-group differences in MAF or BETY-BQ indicates that both interventions carry valuable, yet distinct, therapeutic weights.

The primary limitation of this study is that we did not objectively measure physiological autonomic parameters (e.g.,

heart rate variability, electrodermal activity) or local blood flow changes to directly confirm the mechanisms of action of OMTT. Therefore, our physiological interpretations regarding sympathetic downregulation and parasympathetic activation remain theoretical and speculative, based on indirect behavioral and symptomatic improvements. Additionally, the lack of biochemical markers (e.g., substance P or cortisol levels) limits our ability to objectively map neuroendocrine adaptations.

A key limitation is the small final sample size. Although initial randomization met the power threshold (n=28), dropouts due to personal reasons reduced the analyzed sample to 20. Consequently, the study may be underpowered, increasing Type II error risks or overestimating effect sizes. Furthermore, post-attrition unequal group sizes (n=11 vs. n=9) led to minor, non-significant baseline clinical differences. Although delta change scores minimized their impact, baseline imbalances might have subtly influenced outcomes. Thus, these pilot findings require cautious interpretation and validation through larger, fully powered trials employing restricted block randomization to ensure group equivalence.

Another significant limitation of this study is the lack of blinding for both the participants and the therapist delivering the interventions. Due to the nature of manual therapy and physical exercises, full double-blinding is methodologically challenging. Consequently, this lack of blinding may have introduced performance and expectancy biases, particularly regarding subjective, self-reported outcomes such as pain distribution (WPI), symptom severity (SSS), and sleep quality (PSQI). Participants receiving a more complex, multimodal intervention (OMTT+CP) might have harbored higher expectations of clinical benefit compared to the control group. To mitigate this potential bias, the outcome assessments were strictly conducted by an independent, fully blinded researcher. Furthermore, we included an objective physical metric—PPT via digital algometry—to balance the subjective nature of the questionnaires and to provide a more resilient clinical evaluation.

In conclusion, combining OMTT with CP may yields enhanced clinical outcomes in pain reduction, PPTs, and sleep quality compared to Pilates alone within this pilot sample. Based on these preliminary findings, future multidisciplinary FM rehabilitation programs could consider exploring the integration of OMTT to address autonomic dysfunctions comprehensively, pending confirmation from larger, fully-powered trials.

## CONCLUSION

In conclusion, this pilot study provides encouraging preliminary evidence that osteopathic manual treatment techniques may potentially enhance the effectiveness of exercise in reducing pain and increasing the PPT in individuals with FM. As a preliminary investigation, these findings demonstrate the feasibility of the protocol but cannot be considered definitive. Robust, larger-scale randomized controlled trials with long-term follow-ups are strictly required to verify these clinical trends and establish standard guidelines.

## ETHICAL APPROVAL

Approval for the study was granted by the Local Ethics Committee (decision no: 09, dated: 05.30.2023). Regarding the Clinical Trial Registration, we would like to clarify that this study was conducted as a pilot investigation with a specific focus on evaluating the feasibility and preliminary clinical outcomes of osteopathic manual therapy in fibromyalgia patients. While the study was not prospectively registered in an international clinical trial registry (such as ClinicalTrials.gov), it was performed in strict accordance with the Declaration of Helsinki. Furthermore, the study protocol received full approval from the Local Ethics Committee prior to commencement (Decision No: 09, Date: 05.30.2023).

## AUTHORS' CONTRIBUTIONS

**AD:** Conceptualization, Data curation, Formal Analysis, Writing – original draft, Writing – review & editing. **EGK:** Conceptualization, Data curation, Formal Analysis, Writing – original draft, Writing – review & editing. **GT:** Conceptualization, Data curation, Formal Analysis, Writing – original draft, Writing – review & editing. **BBC:** Conceptualization, Data curation, Formal Analysis, Writing – original draft, Writing – review & editing. **UK:** Conceptualization, Data curation, Formal Analysis, Writing – original draft, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES

1. Smith HS, Harris RE, Clauw DJ. Fibromyalgia. In: Benzon HT, Raja SN, Molloy RE, Liu SS, Fishman SM, editors. Essentials of pain medicine. 3rd ed. Philadelphia: Elsevier; 2011. p. 345-50.
2. Bordoni B, Escher AR, Toccafondi A. Osteopathic manual treatment and fibromyalgia: a possible strategy? Cureus. 2022;14(10):e30276.
3. Calsius J, Bierma-Zeinstra S, Putte L, Graaf Y, Bie R, Ostelo R, et al. A touch of relief: a randomized controlled trial on the effect

of osteopathic treatment on pain and heart rate variability in fibromyalgia patients. J Bodyw Mov Ther. 2021;26:214-20.

4. Nicholas AS, Nicholas EA. Atlas of osteopathic techniques. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2012. p. 86.
5. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken). 2010;62(5):600-10.
6. Okifuji A, Turk DC, Sinclair JD, Starz TW, Marcus DA. A standardized manual tender point survey: I. Development and determination of a threshold point for the identification of positive tender points in fibromyalgia syndrome. J Rheumatol. 1997;24(2):377-83.
7. Salaffi F, Carlo M, Farah S, Marotto D, Atzeni F, Sarzi-Puttini P. The Fibromyalgia Impact Questionnaire (FIQ) and its derived versions: a systematic review of their psychometric properties. Expert Rev Pharmacoecon Outcomes Res. 2021;21(3):345-56.
8. Sezgin H, Erden S. Sleep quality and quality of life in patients with fibromyalgia. Cureus. 2023;15(4):e36611.
9. Smedslund G, Sexton J, Kjeken I, Hagen KB, Dagfinrud H. Fatigue in patients with fibromyalgia: a systematic review and meta-analysis. Musculoskeletal Care. 2023;21(1):1-14.
10. Sanchez-Pucero AJ, Zamorano-San-Segundo MR, de-la-Cruz-Torres B, Barrera-Garcimartin A, Diaz-Maroto A, Albornoz-Cabello M. Quality of life in women with fibromyalgia: a cross-sectional study. Int J Environ Res Public Health. 2022;19(18):11425.
11. Koca TT, Seyithanoglu M, Boyaci A, Arkan H. Examination of exercise benefits/barriers scale and its relationship with disease activity in fibromyalgia syndrome. Turk J Phys Med Rehabil. 2021;67(2):212-9.
12. Simsir G, Kaymak C. Anxiety and depression levels in patients with fibromyalgia during the COVID-19 pandemic. J Clin Med. 2024;13(4):985.
13. Bakar Y, Tugral A, Atasavun Uysal S, Ozyurek S, Cavlak U. Psychometric properties of the Turkish version of the BETY-Biopsychosocial Questionnaire in patients with fibromyalgia. Rheumatol Int. 2022;42(10):1833-41.
14. Albers J, Jäkel A, Wellmann K, Hehn U, Schmidt T. Effectiveness of 2 osteopathic treatment approaches on pain, pressure-pain threshold, and disease severity in patients with fibromyalgia: a randomized controlled trial. Complement Med Res (Basel). 2018;25(2):122-8.
15. Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Häuser W, Fluss E, et al. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis. 2017;76(2):318-28.
16. Schaffer AL, Schaffer CL, Ferreira J, Silva S, Paula R. The effectiveness of pilates exercise in people with fibromyalgia: a systematic review and meta-analysis. Clin Rehabil. 2022;36(11):1413-28.
17. Sirianni J, Ibrahim M, Patwardhan A. Chronic pain syndromes, mechanisms, and current treatments. Prog Mol Biol Transl Sci. 2015;131:565-611.
18. Reyes Paso GA, Coba P. Reduced activity, reactivity and functionality of the sympathetic nervous system in fibromyalgia: an electrodermal study. PLoS One. 2020;15(10):e0241154.
19. Katz DL, Greene L, Ali A, Faridi Z. The pain of fibromyalgia syndrome is due to muscle hypoperfusion induced by regional vasomotor dysregulation. Med Hypotheses. 2007;69(3):517-25.
20. Cerritelli F, Cardone D, Pirino A, Merla A, Scoppa F. Does osteopathic manipulative treatment induce autonomic changes in healthy participants? A thermal imaging study. Front Neurosci. 2020;14:887.
21. Schleip R, Findley TW, Chaitow L, Huijing PA, editors. Fascia: the tensional network of the human body. Edinburgh: Churchill Livingstone Elsevier; 2012.
22. Arienti C, Bosisio T, Ratti S, Miglioli R, Negrini S. Osteopathic manipulative treatment in women with fibromyalgia: a randomized controlled clinical trial. Front Med (Lausanne). 2022;9:855513.
23. Dal Farra F, Chiesa A, Risio RG, Vismara L, Bergna A. Fast improvements in functional status after osteopathic manipulative treatment based on myofascial release in patients with moderate or severe fibromyalgia: a retrospective study. J Complement Integr Med. 2021;19(2):481-7.
24. Coste J, Medkour T, Maigne JY, Dubourthoumieu N, Bert F. Osteopathic medicine for fibromyalgia: a sham-controlled randomized clinical trial. Ther Adv Musculoskelet Dis. 2021;13:1759720X211009017.
25. Schulze NB, Salemi MM, Alencar GG, Moreira MC, Siqueira GR. Efficacy of manual therapy on pain, impact of disease, and quality of life in the treatment of fibromyalgia: a systematic review. Pain Physician. 2020;23(5):461-76.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260662

# Post-cesarean analgesia: transversalis fascia plane block versus wound infiltration in a randomized trial

Gülşen Cebecik Teomete[1], Ecem Güçlü Öztürk[2], Irmak Çimenoğlu[2], Merve Ergenç[2], Beliz Bilgili[2]*

## SUMMARY

OBJECTIVE: The aim of this study was to compare ultrasound-guided transversalis fascia plane block with local wound infiltration using equivalent local anesthetic doses within a standardized multimodal analgesia protocol.

METHODS: This prospective, single-blind randomized trial was conducted between May and June 2024 at a tertiary care hospital. Women with American Society of Anesthesiologists physical status II–III undergoing cesarean delivery under spinal anesthesia received either bilateral ultrasound-guided transversalis fascia plane block or local wound infiltration with the same total dose and volume of bupivacaine. All patients were managed with standardized multimodal analgesia. Patients with contraindications to the study interventions or inability to use patient-controlled analgesia were excluded. The primary outcome was 24-h opioid consumption. Secondary outcomes included pain scores, rescue analgesia use, opioid-related adverse events, and complications.

RESULTS: Forty-two patients (n=21/group) were analyzed. Twenty-four-hour opioid consumption was similar between transversalis fascia plane block and local wound infiltration groups (median 100 [65–152] mg vs. 85 [65–135] mg; p=0.554). Pain scores remained low and comparable throughout 48 h. Rescue analgesia was rarely required, and no postoperative nausea, opioid-related adverse events, or procedure-related complications were observed.

CONCLUSION: In a standardized multimodal analgesia setting after cesarean delivery, transversalis fascia plane block did not reduce opioid consumption or improve postoperative pain scores compared with local wound infiltration.

KEYWORDS: Cesarean section. Pain, postoperative. Nerve block. Ultrasonography. Analgesics, opioid.

## INTRODUCTION

According to World Health Organization (WHO) data, cesarean section (CS) rates are increasing, currently at 21% of births and projected to reach 29% by 2030[1]. Effective postoperative analgesia is essential, particularly as cesarean section (CS) requires techniques that are safe for both mother and newborn[2]. In obstetric practice, simple and reproducible methods are especially valuable in high-volume settings. The PROcedure-SPECific Postoperative Pain ManagemenT (PROSPECT) guideline recommends local wound infiltration and/or regional techniques (e.g., transversus abdominis or quadratus lumborum blocks) with Grade A evidence for pain control[3]. A meta-analysis reported reductions in 24-h morphine consumption of 11 mg with transversalis fascia plane block (TFPB) and 8 mg with local wound infiltration (LWI)[4].

The TFPB is a regional technique targeting the iliohypogastric and ilioinguinal nerves in their proximal course and has been proposed as an effective option for lower abdominal surgery[5,6]. Although several trials[7,8] have evaluated transversalis fascia plane block for cesarean analgesia, reported mixed results and many comparisons were limited by heterogeneous postoperative analgesic regimens or unequal local anesthetic dosing between study groups. Consequently, whether this technically more complex block provides meaningful benefit over simpler wound infiltration within contemporary multimodal analgesia pathways remains uncertain.

This randomized controlled trial compared ultrasound-guided transversalis fascia plane block with local wound infiltration after cesarean delivery, using an equivalent total dose of local anesthetic within a standardized multimodal analgesia protocol. We hypothesized that TFPB would reduce 24-h postoperative opioid consumption compared with LWI. The primary outcome was total opioid consumption in the first 24 h. Secondary outcomes included Numeric Rating Scale (NRS) scores, rescue analgesia requirements, and postoperative nausea and vomiting (PONV).

[1]Marmara University, Pendik Training and Research Hospital, Department of Anesthesiology and Reanimation – Istanbul, Turkey.
[2]Marmara University, School of Medicine, Department of Anesthesiology and Reanimation – Istanbul, Turkey.
*Corresponding author: beliz.bilgili@marmara.edu.tr
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on April 21, 2026. Accepted on July 24, 2026.
Scientific Editor: José Maria Soares Júnior

## METHODS

This prospective, single-blinded, randomized controlled trial was approved by the Marmara University Faculty of Medicine Ethics Committee (Approval Date: 14.05.2024, No: 09.2024.217) and registered at clinicaltrials.gov (NCT06425718) before enrollment. Written informed consent was obtained from all participants, and the study adhered to the Declaration of Helsinki. Between May and June 2024, consecutive patients meeting the eligibility criteria who were scheduled to undergo cesarean section under spinal anesthesia with American Society of Anesthesiologists (ASA) physical status II–III were invited to participate. Exclusion criteria included: refusal or withdrawal of consent, inability to give consent, allergy to local anesthetics, block area infection, chronic opioid use and inability to use the patient-controlled analgesia (PCA) system.

Randomization and blinding: Participants were randomized using a computer-generated sequence (randomizer.org). Allocation was concealed with sequentially numbered, opaque, sealed envelopes opened immediately before intervention. All blocks were performed by a single experienced anesthesiologist not involved in data collection or analysis, while LWI was performed by the surgeon. Postoperative outcomes were assessed by a blinded pain team, and patients were also blinded, with interventions concealed under standard dressings. The study followed Consolidated Standards of Reporting Trials (CONSORT) guidelines (Figure 1).

Intraoperative management: All patients received standardized anesthesia. Standard monitoring [electrocardiography (ECG), noninvasive blood pressure, pulse oximetry] was applied, with oxygen (4 L/min) via face mask and a 20 G intravenous (IV) cannula. Patients <150 or >150 cm received 1.8 or 2.0 mL hyperbaric bupivacaine (5 mg/mL) with 20 mcg fentanyl intrathecally, respectively. Ondansetron (4 mg IV) was given for antiemetic prophylaxis, and oxytocin was administered post-delivery per protocol.

At surgery end, in the TFPB group, a curvilinear probe (2–5 MHz Sparq Ultrasound; Philips, USA) was placed in the midaxillary line between the iliac crest and rib cage. After identifying the external oblique, internal oblique, and transversus abdominis muscles, 20 mL of 0.25% bupivacaine was injected bilaterally using an in-plane technique following saline confirmation. In the LWI group, 40 mL of 0.25% bupivacaine was infiltrated subcutaneously into the wound by the surgeon. Both groups received IV paracetamol (1 g), IV dexamethasone

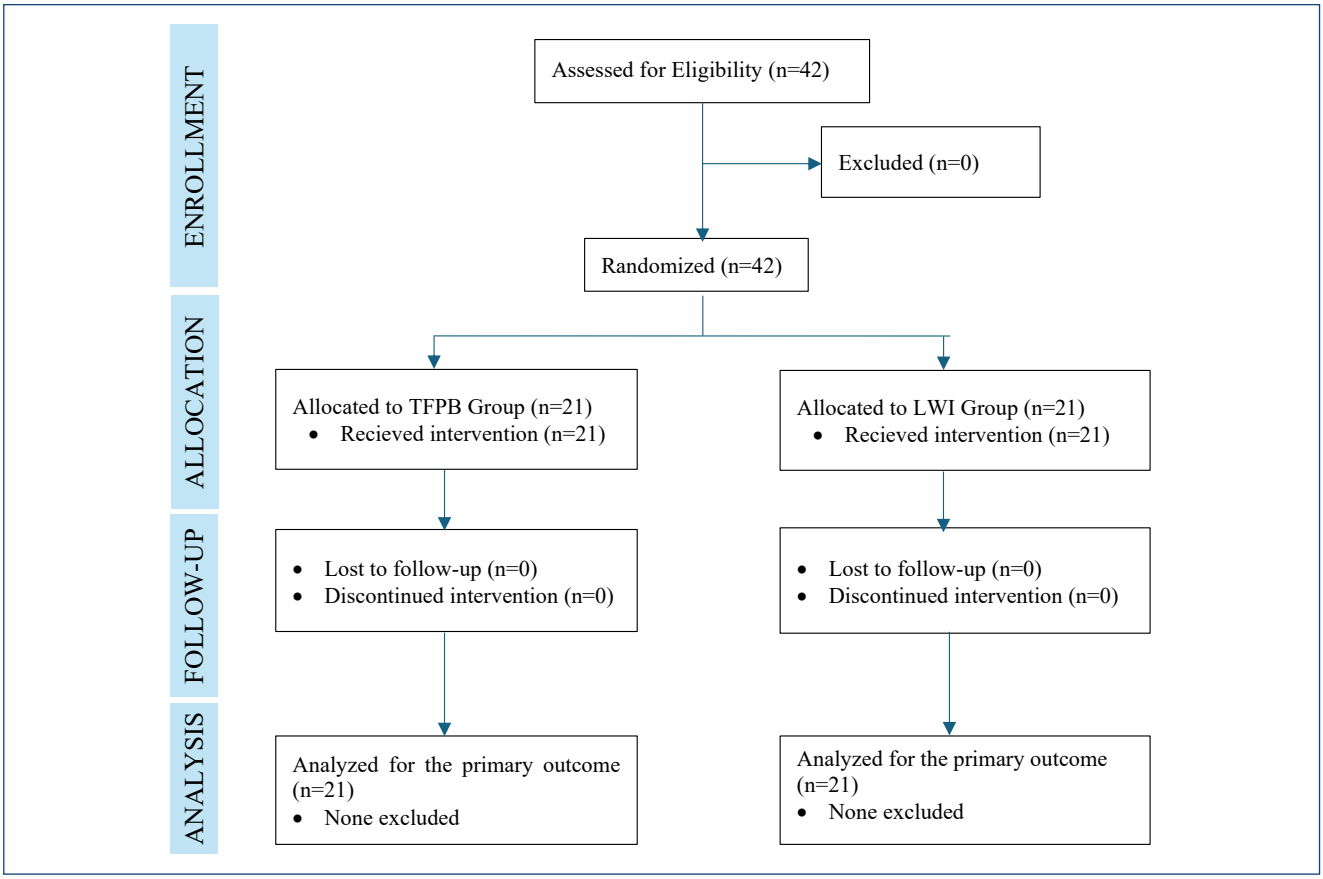



**Figure 1.** Consolidated Standards of Reporting Trials flow diagram of the study. TFPB: transversalis fascia plane block; LWI: local wound infiltration.

(8 mg), and intramuscular (IM) diclofenac (75 mg) for multimodal analgesia.

Postoperative management: All patients received IV paracetamol (1 g every 8 h) and IV patient-controlled analgesia with tramadol (5 mg/mL; no basal infusion, 1 mL bolus, 10 min lockout). If NRS >3, IM diclofenac (75 mg; max two doses/day) was administered as rescue analgesia.

In the TFPB group, sensory block was assessed after recovery from spinal anesthesia using cold and pinprick testing to confirm dermatomal involvement and exclude block failure, defined as absence of sensory changes. The block was evaluated by the anesthesiologist performing TFPB, who was not involved in data collection or analysis. Block duration was not assessed, as comparison with local wound infiltration was not feasible.

Data collection: Baseline demographic and intraoperative data were recorded. Postoperative assessments were performed at 0, 6, 12, 24, and 48 h by a pain team blinded to group allocation. At each assessment, NRS score, total tramadol consumption, rescue analgesia requirement and opioid-related adverse events were recorded.

Outcome measures: The primary outcome was 24-h total tramadol consumption. Secondary outcomes included NRS scores (0–10), rescue analgesia requirements (defined as NRS >3), and opioid-related adverse events (nausea/vomiting, urinary retention, sedation, pruritus, and respiratory depression) recorded as present or absent.

Sample size: Pilot data for sample size estimation were retrospectively obtained from institutional records of CS patients managed with the same analgesic protocol. The pilot dataset included 12 patients per group and was used solely for variance estimation and effect size approximation. The sample size was estimated using pilot data in which mean (standard deviation [SD]) tramadol consumption at 24 h postoperatively was 135.0 (51.3) mg in the control group and 85.4 (49.5) mg in the intervention group (≈37% reduction). Using G*Power 3.1 for a two-tailed Wilcoxon-Mann-Whitney test ($\alpha$=0.05, power=0.80; Vargha-Delaney A≈0.76), 19 patients per group were required[9]. Allowing for 10% dropout, 21 patients per group were planned (total n=42).

Statistical analysis: Statistical analysis was performed using IBM Statistical Package for the Social Sciences (IBM SPSS) Statistics v27. Data distribution was assessed with the Shapiro-Wilk test and histograms. Continuous variables are presented as mean±SD or median [interquartile range (IQR)], as appropriate. The primary outcome (24-h opioid consumption) was non-normally distributed and analyzed using the Mann-Whitney U test with confidence intervals. Secondary continuous outcomes were analyzed similarly, while categorical variables were compared using Fisher's exact test. All tests were two-tailed, with $p<0.05$ considered significant.

## RESULTS

Forty-two patients were randomized and completed the study. We did not observe any block failure. The CONSORT flow diagram is shown in Figure 1.

Baseline demographic and clinical characteristics were comparable between the two groups, with no statistically significant differences in any baseline variables (Table 1). The primary outcome, total opioid consumption during the first 24 postoperative hours, did not differ significantly between the TFPB and LWI groups (median 100 [65–152] mg vs. 85 [65–135] mg, 95%CI -30.12, 51.27, p=0.554). Median opioid consumptions at all time intervals were comparable among the study groups (Table 2).

Postoperative pain scores remained low and were comparable between groups throughout the 48-h follow-up period (Table 3). Median pain scores were recorded to be <3/10 across all time intervals, indicating that postoperative pain was generally well controlled within the target range. Rescue analgesia was uncommon. In the LWI group, three patients (14.2%) required rescue analgesia, whereas no patients in the TFPB group required rescue analgesia; however, this difference was not statistically significant (p=0.072).

**Table 1.** Demographic data and characteristics of the participants.

| | | Total (n=42) | Group TFPB (n=21) | Group LWI (n=21) | p-value | Std diff |
|---|---|---|---|---|---|---|
| Age (year) | | 31.5±6.5 | 32.5±6.0 | 30.4±6.9 | 0.319[a] | 0.32 |
| BMI (kg/m$^2$) | | 30.7±4.4 | 29.9±3.9 | 31.4±4.9 | 0.505[a] | 0.34 |
| ASA | II | 40 | 21 (100) | 19 (90) | 0.152[b] | |
| | III | 2 | 0 (0) | 2 (10) | | |

[a]Mann-Whitney-U test. [b]Fisher's exact test. Values are presented as number of patients (%) or mean±SD. ASA: American Society of Anesthesiologists physical status; BMI: body mass index; TFPB: transversalis fascia plane block; LWI: local wound infiltration. Standardized differences are presented for continuous variables only.

**Table 2.** Comparison of total tramadol consumption during predefined postoperative time intervals between the transversalis fascia plane block and local wound infiltration groups.

| IV PCA | | Total (n=42) | Group | | 95%CI of the difference | p-value[a] |
|---|---|---|---|---|---|---|
| | | | Group TFPB (n=21) | Group LWI (n=21) | | |
| 0–6th hour | Median (Q1, Q3) | 30 (20, 61.25) | 45 (20, 70) | 25 (17.5, 50) | -7.67, 35.77 | 0.193 |
| 6–12th hour | Median (Q1, Q3) | 30 (15, 50) | 25 (12.5, 55) | 30 (20, 42.5) | -22.51, 15.37 | 0.528 |
| 12–24th hour | Median (Q1, Q3) | 30 (20, 36.25) | 30 (17.5, 37.5) | 30 (22.5, 37.5) | -12.57, 12.57 | 0.674 |
| 24–48th hour | Median (Q1, Q3) | 50 (35, 70) | 40 (32.5, 75) | 50 (45, 75) | -25.88, 15.88 | 0.332 |
| 24th hour total | Median (Q1, Q3) | 90 (65, 142.5) | 100 (65, 152) | 85 (65, 135) | -30.12, 51.27 | 0.554 |

Summary statistics are reported as median (interquartile range). Postoperative opioid consumption is expressed in milligrams (mg). CI: confidence interval; IV PCA: intravenous patient-controlled analgesia; NRS: Numeric Rating Scale; TFPB: transversalis fascia plane block; LWI: local wound infiltration. [a]Compared by Mann-Whitney-U test.

**Table 3.** Postoperative pain scores (Numeric Rating Scale) at predefined postoperative time points in the transversalis fascia plane block and local wound infiltration groups.

| NRS | | Total (n=42) | Group | | 95%CI of the difference | p-value[a] |
|---|---|---|---|---|---|---|
| | | | Group TFPB (n=21) | Group LWI (n=21) | | |
| 6th hour | Median (Q1, Q3) | 2 (1, 2) | 2 (1, 3) | 2 (1, 2) | -0.56, 0.94 | 0.70 |
| 12th hour | Median (Q1, Q3) | 1 (1, 2) | 1 (1, 2) | 2 (1, 2) | -0.89, 0.22 | 0.20 |
| 24th hour | Median (Q1, Q3) | 1 (1, 2) | 1 (1, 2) | 2 (1, 2) | -0.84, 0.27 | 0.28 |
| 48th hour | Median (Q1, Q3) | 1 (1, 1) | 1 (0, 1) | 1 (1, 1.5) | -0.67, 0.9 | 0.14 |

CI: confidence interval; NRS: Numeric Rating Scale; TFPB: transversalis fascia plane block; LWI: local wound infiltration. [a]Compared by Mann-Whitney-U test.

PONV requiring medication was not observed in either group. Furthermore, no patients experienced less common opioid-related side effects such as itching, respiratory depression, or urinary retention. Additionally, no complications related to TFPB or LWI were observed.

## DISCUSSION

In this randomized trial evaluating postoperative analgesia after cesarean delivery, ultrasound-guided transversalis fascia plane block did not reduce 24-h opioid consumption compared with local wound infiltration within a multimodal analgesia protocol. Pain scores, rescue analgesia use, and opioid-related adverse events were similar. These findings suggest no clinically meaningful advantage of the more complex fascial plane block over wound infiltration.

TFPB is a deep abdominal wall block targeting the proximal ilioinguinal and iliohypogastric nerves, potentially providing broader dermatomal coverage than superficial blocks[5,10,11]. Injection beneath the transversalis fascia may allow spread toward thoracolumbar nerves before lateral branching, theoretically enhancing analgesia in lower abdominal surgery[12-14]. However, this trial showed no clinically meaningful opioid-sparing effect of TFPB when compared with LWI. Several factors may explain the lack of a statistically significant difference. Both interventions were applied within a standardized multimodal regimen including spinal anesthesia, regular paracetamol, nonsteroidal anti-inflammatory drug (NSAID) rescue analgesia, and tramadol PCA. Contemporary cesarean analgesia increasingly depends on multimodal regimens that maintain acceptable pain control while reducing opioid requirements. In this setting, baseline analgesia may already provide adequate pain relief, thereby diminishing observable differences between regional techniques. Consequently, the additional analgesic benefit of further fascial plane blocks within such pathways may be limited.

Previous studies on TFPB in cesarean delivery show heterogenous results. Compared with inactive or sham controls, TFPB reduces postoperative opioid consumption, supporting its analgesic efficacy[12,13]. However, its advantage appears less pronounced when compared with other active analgesic techniques. In a comprehensive meta-analysis on the effectiveness of local anesthetic techniques for post-CS analgesia, it was found that wound infiltration reduced total morphine consumption at hours 12 and 24 compared to inactive controls, while TFPB showed similar results to the inactive controls[14]. The only randomized trial directly comparing TFPB with wound infiltration

found similar analgesic requirements, consistent with the present study[15]. Chilkoti et al. reported longer time to first analgesic request and lower NRS scores with TFPB at some time points, but pain scores remained <3 and opioid use was minimal despite the absence of adjuncts in the subarachnoid block, limiting clinical relevance. In addition, unequal local anesthetic dosing (20 vs. 40 mL, both 0.375% ropivacaine) may have biased results in favor of TFPB. In contrast, the present study standardized both volume and dose (40 mL of 0.25% bupivacaine in each group), reducing potential bias and enabling a more balanced comparison of the two techniques.

Evidence from other surgical procedures supports this interpretation. In patients undergoing iliac crest bone graft harvesting, TFPB reduced opioid consumption during the first eight postoperative hours compared with a sham block, although this benefit was not sustained thereafter[16]. In another study of inguinal hernia repair, quadratus lumborum block was associated with lower postoperative opioid consumption than TFPB[17]. These findings suggest that TFPB provides effective postoperative analgesia, but its opioid-sparing effect is more evident when compared with placebo or no regional block than with another effective regional analgesic technique.

Within an optimized multimodal cesarean analgesia pathway and with standardized local anesthetic dosing, our findings suggest that the additional benefit of complex regional techniques is limited, and LWI provides comparable analgesia with greater practicality. LWI is simpler than TFPB, requiring no ultrasound guidance or advanced expertise, has a shorter learning curve, and can be performed by the surgeon during closure. This increases its feasibility, particularly in resource-limited settings. Overall, these results support LWI as a pragmatic option within ERAS protocols, aligning with goals of reduced opioid use, early mobilization, and enhanced recovery[18].

We observed no postoperative nausea and vomiting in either group. The traditional belief that the use of opioids contributes to nausea and vomiting has been recently questioned, as low-dose lipophilic opioids (e.g., fentanyl) used as spinal adjuvants may have antiemetic effects[19,20]. No TFPB-related complications (organ/vascular injury, infection, motor weakness) were observed. As TFPB is performed more posteriorly than transversus abdominis plane (TAP) block, and the perinephric fat lies beneath the transversalis fascia, the risk of peritoneal puncture is minimal.

A key strength of this trial is the standardization of total local anesthetic dose and volume between groups, minimizing a potential source of bias seen in previous studies where unequal dosing may have influenced outcomes. In addition, postoperative analgesia was standardized, allowing a more balanced comparison of techniques. All block procedures were performed by a single experienced anesthesiologist, reducing inter-operator variability.

The study has several limitations. The sample size, based on pilot data, may have overestimated the effect size; however, differences in opioid use were minimal and unlikely to be clinically meaningful. The lack of statistical significance should be interpreted within the context of optimized multimodal analgesia, where incremental benefits may be limited, rather than attributed solely to insufficient power. The absence of a placebo/sham group prevents assessment of absolute efficacy. The anesthesiologist who performed the TFPB also assessed block success, introducing a potential source of assessment bias, although postoperative outcomes were evaluated by blinded investigators. Finally, block duration was not evaluated beyond confirmation of success.

Future studies with larger cohorts and a sham-controlled design are needed to confirm these findings. Research should also evaluate adjuncts such as dexamethasone or dexmedetomidine to enhance the opioid-sparing effect of TFPB, while assessing their safety in postpartum patients, particularly regarding neonatal exposure via breastmilk. In addition, patient-centered outcomes, including mobility, breastfeeding comfort, and quality of recovery, should be incorporated to better define the clinical value of regional techniques in obstetric anesthesia.

## CONCLUSION

In conclusion, within a standardized multimodal cesarean analgesia protocol, ultrasound-guided TFPB did not reduce postoperative opioid consumption compared with LWI. With standardized local anesthetic dosing, the additional complexity of TFPB did not confer a clinically meaningful advantage. Both techniques provided comparable analgesia with low pain scores. Given its simplicity and practicality, LWI remains a reasonable option for routine use.

## AUTHORS' CONTRIBUTIONS

**GCT:** Conceptualization, Data curation. **EGÖ:** Data curation, Formal Analysis. **IÇ:** Investigation, Writing – original draft. **ME:** Investigation, Writing – original draft. **BB:** Conceptualization, Methodology, Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Roberts CL, Nippita TA. International caesarean section rates: the rising tide. Lancet Glob Health. 2015;3(5):e241-2. https://doi.org/10.1016/S2214-109X(15)70111-7
2. Pan PH. Post cesarean delivery pain management: multimodal approach. Int J Obstet Anesth. 2006;15(3):185-8. https://doi.org/10.1016/j.ijoa.2006.04.004
3. Roofthooft E, Joshi GP, Rawal N, Velde M, Joshi GP, Pogatzki-Zahn E, et al. PROSPECT guideline for elective caesarean section: updated systematic review and procedure-specific postoperative pain management recommendations. Anaesthesia. 2021;76(5):665-80. https://doi.org/10.1111/anae.15339
4. Singh NP, Monks D, Makkar JK, Palanisamy A, Sultan P, Singh PM. Efficacy of regional blocks or local anaesthetic infiltration for analgesia after caesarean delivery: a network meta-analysis of randomised controlled trials. Anaesthesia. 2022;77(4):463-74. https://doi.org/10.1111/anae.15645
5. Vasques F, Stecco C, Mitri R, Caro R, Fusco P, Behr AU. Blocking around the transversalis fascia: behind the scene. Minerva Anestesiol. 2019;85(1):15-20. https://doi.org/10.23736/S0375-9393.18.12479-5
6. Singh S, Singh S, Sharma M. Successful use of transversalis fascia plane block for iliac crest bone harvesting in a polytrauma patient: a case report. Saudi J Anaesth. 2020;14(1):107-8. https://doi.org/10.4103/sja.SJA_403_19
7. Bilgin S, Aygün H, Genç C, Dost B, Tulgar S, Kaya C, et al. Comparison of ultrasound-guided transversalis fascia plane block and anterior quadratus lumborum block in patients undergoing caesarean delivery: a randomized study. BMC Anesthesiol. 2023;23(1):249. https://doi.org/10.1186/s12871-023-02206-w
8. Dost B, Cassai A, Bugada D, Balzani E, Karapınar YE, Beldaglı M, et al. The analgesic effect of transversalis fascia plane block after cesarean delivery: a systematic review and meta-analysis with trial sequential analysis. Minerva Anestesiol. 2025;91(6):535-45. https://doi.org/10.23736/S0375-9393.25.18771-3
9. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods. 2007;39(2):175-91. https://doi.org/10.3758/bf03193146
10. Nair A, Dudhedia U, Rangaiah M, Borkar N. Ultrasound-guided transversalis fascia plane block for postoperative analgesia: a systematic review and meta-analysis. Indian J Anaesth. 2023;67(4):331-42. https://doi.org/10.4103/ija.ija_43_23
11. Hebbard PD. Transversalis fascia plane block, a novel ultrasound-guided abdominal wall nerve block. Can J Anaesth. 2009;56(8):618-20. https://doi.org/10.1007/s12630-009-9110-1
12. Aydin ME, Bedir Z, Yayik AM, Celik EC, Ates I, Ahiskalioglu EO, et al. Subarachnoid block and ultrasound-guided transversalis fascia plane block for caesarean section: a randomised, double-blind, placebo-controlled trial. Eur J Anaesthesiol. 2020;37(9):765-72. https://doi.org/10.1097/EJA.0000000000001222
13. Serifsoy TE, Tulgar S, Selvi O, Senturk O, Ilter E, Peker BH, et al. Evaluation of ultrasound-guided transversalis fascia plane block for postoperative analgesia in cesarean section: a prospective, randomized, controlled clinical trial. J Clin Anesth. 2020;59:56-60. https://doi.org/10.1016/j.jclinane.2019.06.025
14. Wang J, Zhao G, Song G, Liu J. The efficacy and safety of local anesthetic techniques for postoperative analgesia after cesarean section: a Bayesian network meta-analysis of randomized controlled trials. J Pain Res. 2021;14:1559-72. https://doi.org/10.2147/JPR.S313972
15. Chilkoti GT, Gaur D, Saxena AK, Gupta A, Agarwal R, Jain S. Ultrasound-guided transversalis fascia plane block versus wound infiltration for both acute and chronic post-caesarean pain management: a randomised controlled trial. Indian J Anaesth. 2022;66(7):517-22. https://doi.org/10.4103/ija.ija_173_22
16. Black ND, Malhas L, Jin R, Bhatia A, Chan VWS, Chin KJ. The analgesic efficacy of the transversalis fascia plane block in iliac crest bone graft harvesting: a randomized controlled trial. Reg Anesth Pain Med. 2022;47(8):499-504.
17. Naveenkumar P, Pandian N, Shanmugasundaram S, Kumar BA, Periasamy P. Comparative efficacy of ultrasound-guided transversalis fascia plane block versus quadratus lumborum block for postoperative analgesia in inguinal hernia repair. Turk J Anaesthesiol Reanim. 2024;52(4):199-205.
18. Macones GA, Caughey AB, Wood SL, Wrench IJ, Huang J, Norman M, et al. Guidelines for postoperative care in cesarean delivery: Enhanced Recovery After Surgery (ERAS) Society recommendations (part 3). Am J Obstet Gynecol. 2019;221(3):247.e1-9. https://doi.org/10.1016/j.ajog.2019.04.012
19. Swain A, Nag DS, Sahu S, Samaddar DP. Adjuvants to local anesthetics: current understanding and future trends. World J Clin Cases. 2017;5(8):307-23. https://doi.org/10.12998/wjcc.v5.i8.307
20. Caballero-Lozada AF, Gómez JM, Torres-Mosquera A, González-Carvajal Á, Marín-Prado A, Zorrilla-Vaca A, et al. Impacts of intrathecal fentanyl on the incidence of postoperative nausea/vomiting: systematic review and meta-analysis of randomized studies. J Anaesthesiol Clin Pharmacol. 2022;38(4):529-36. https://doi.org/10.4103/0970-9185.369222

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260706

# Retrospective study of inflammatory markers in high-grade cervical lesions

Pınar Yıldız[1]*, Güneş Topçu[2], Gerçek Aydın[3], Asiye Uzun[2]

## SUMMARY

**OBJECTIVE:** The aim of this study was to evaluate the predictive value of systemic inflammation markers for abnormal cervical screening results and high-grade cervical lesions in women participating in a human papillomavirus- and cytology-based screening program, thereby determining colposcopy indications more accurately to reduce unnecessary colposcopy requests and prevent overtreatment.

**METHODS:** A total of 2,758 women who underwent cervical cancer screening between 2020 and 2025 were included in this retrospective cross-sectional study. Abnormal screening results were defined as abnormal cytology (≥Atypical squamous cells of undetermined significance [ASC-US]) and/or high-risk human papillomavirus positivity. Hematological parameters measured within 1 month prior to screening were used to calculate neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, and systemic immune-inflammation index. Continuous variables were evaluated using nonparametric tests, and the independent predictors of abnormal screening results and high-grade lesions were identified using logistic regression models.

**RESULTS:** Neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, and systemic immune-inflammation index values were significantly higher in the abnormal smear groups and in human papillomavirus -positive patients (all $p<0.001$). Neutrophil-to-lymphocyte ratio displayed the highest diagnostic performance for high-grade lesions (area under the curve=0.871; cutoff value >4.41, sensitivity 66.7%, specificity 99.2%), followed by systemic immune-inflammation index (area under the curve=0.847), while platelet-to-lymphocyte ratio showed lower specificity (area under the curve=0.741).

**CONCLUSION:** In conclusion, our findings suggest that inflammatory markers are associated with abnormal cervical screening results. To determine the significance and clinical utility of serum inflammatory markers, multicenter clinical trials involving larger patient cohorts are needed.

**KEYWORDS:** Cervical intraepithelial neoplasia. Human papillomavirus. Papanicolaou test. Inflammation.

## INTRODUCTION

Cervical cancer is a type of malignant tumor affecting the female reproductive system; it is the most common gynecological cancer in developing countries and accounts for the fourth most common type of cancer worldwide[1] Global Cancer Statistics (GLOBOCAN 2022) data indicate that 662,301 women were diagnosed with cervical cancer in 2022, and an estimated 348,874 died from the disease[2]. As cancer screening programs have become more widely available, it is known that the rate of diagnosed cervical intraepithelial neoplasia (CIN) has increased. In contrast, the incidence of advanced-stage cancer has decreased[3]. Early detection plays an important role in the prevention and treatment of cervical cancer[4]. During clinical practice, women who receive abnormal screening results should be evaluated with a colposcopy and biopsy to detect precancerous lesions and cervical cancer[5].

Human papillomavirus infection is a major risk factor associated with the development of cancer and can potentially lead to CIN. This infection is relatively common and affects approximately 90% of people during their lifetime, potentially leading to serious health issues. There are more than 200 genotypes of HPV. While the majority of them are not associated with complications and resolve spontaneously, some types are considered high-risk and can lead to cancer[6,7].

In recent studies, the potential value of hematological markers in the early diagnosis and prognosis of cancers has been recognized. It is considered that inflammatory parameters such as the neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and systemic immune-inflammation index (SII) are correlated with tumor cell formation, progression, and prognosis[8]. The cytokines, vascular endothelial growth factor, and chemokines produced by neutrophils are involved in tumor growth, angiogenesis, and metastasis. In the same way, platelets stimulate tumor cell proliferation and help these cells evade immune control, thereby promoting tumor spread[9]. The is a novel marker of inflammation and a useful biomarker for assessing the

[1]Academic Hospital, Department of Obstetrics and Gynecology – Istanbul, Turkey.
[2]Bahçelievler Medipol University Hospital, Department of Obstetrics and Gynecology – Istanbul, Turkey.
[3]Istanbul Health and Technology University, Department of Obstetrics and Gynecology – Istanbul, Turkey.
*Corresponding author: pinar.9530@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on March 12, 2026. Accepted on July 19, 2026.
Scientific Editor: José Maria Soares Júnior

prognosis of various cancers; it is calculated using the formula: platelet count × neutrophil count/lymphocyte count[10,11]. Several inflammatory markers in the blood are linked to the formation and development of tumors in various types of cancer[12,13]. In our study, 2,758 patients who had undergone smear tests were retrospectively analyzed to ensure accurate patient assessment, improve screening effectiveness, prevent unnecessary colposcopy and biopsy referrals, and reduce overtreatment to identify more effective alternative biomarkers for predicting cervical lesions and establishing more appropriate clinical criteria for colposcopy.

## METHODS

### Study design and setting

This study was designed and reported in accordance with the STROBE Statement. It was conducted as a single-center, retrospective descriptive study at the Department of Obstetrics and Gynecology of a tertiary care university hospital in Istanbul, Türkiye.

### Study population and sampling strategy

The study population consisted of women who presented to our outpatient clinic, had a complete blood count within 1 month, and underwent cervical cytology screening between January 1, 2020, and January 1, 2025. During the study period, a total of 2,758 pap smear tests and blood tests were performed. All women aged 21–65 years who were not pregnant, had no history of treatment for cervical lesions, had a pap smear and/or HPV testing and a blood test within the defined period, and had complete medical records were included. Patients under 21 or over 65 years of age, those with organ failure, people with an active infection or malignant neoplastic diseases, those who had been treated with antibiotics before testing, those taking medication impacting inflammatory markers, and pregnant women were excluded. A total of 82 cases were excluded due to not meeting the inclusion criteria and/or having incomplete data.

### Research questions

The study aimed to evaluate effective alternative biomarkers for predicting cervical lesions and to assess their impact on clinical processes in determining the requirement for colposcopy.

### Data collection and variables

Data were retrospectively obtained from the hospital's Electronic Health Record (EHR) system. The following variables were recorded: age, parity, pap smear and HPV test results, PLR, NLR, and SII values based on the complete blood count results. All data were anonymized before analysis.

### Definition of exposure and outcome measures

Exposure variables were defined as hematological inflammation parameters, including NLR, PLR, and SII, derived from complete blood count measurements obtained within 1 month preceding cervical cancer screening.

NLR=neutrophil count/lymphocyte count
PLR=platelet count/lymphocyte count
SII=(platelet × neutrophil)/lymphocyte

We defined the primary outcome as the presence of an abnormal cervical screening result. Abnormal cytology (≥ASC-US) and/or high-risk HPV positivity were considered to constitute this outcome. Secondary outcomes consisted of the presence of high-grade cytological abnormalities (≥HSIL), evaluation of the relationship between inflammatory indicators and the presence of HPV, and investigation of the performance of hematological inflammation indices in the prediction of abnormal screening results.

### Follow-up

No follow-up was conducted because of the study's retrospective design.

### Handling of missing data

Cases with incomplete or missing data were excluded from the study.

### Statistical analysis

Statistical analyses were performed using SPSS version 27. Quantitative variables were expressed as mean, standard deviation, median, minimum, and maximum, while qualitative variables were presented as frequencies and percentages. The Fisher-Freeman-Halton exact test was used to compare qualitative data, and $p<0.01$ was considered statistically significant.

### Ethical considerations

Ethical approval was obtained from the Institutional Ethics Committee (Approval No: 5163; Date: November 11, 2025). The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and the ethical standards of the country of origin. The authors declare no conflicts of interest.

## RESULTS

A total of 2,758 patients were included in the study. Continuous variables deviated from a normal distribution, as expected given the large sample size (Shapiro-Wilk, $p<0.001$), and were

reported as medians (IQR); categorical variables were presented as frequencies and percentages.

The median age of the study population was 37 [30–42] years, and 61.2% of the women (n=1,689) were multiparous. Based on the smear results, 66.8% of patients (n=1,841) had benign cytology, 18.2% (n=503) were ASC-US, and 11.3% (n=312) were LSIL. High-grade lesions (ASC-H, HSIL, AGC, AIS, and SCC/AdenoCa) accounted for 3.7% (n=102) of all cases. HPV positivity was observed in 20.8% of patients (n=573), and the most common subtype was HPV 16 (7.7%) (Table 1).

NLR, PLR, and SII values were significantly increased in line with cytological grade and were higher in groups with abnormal smear results compared to benign cases (all p<0.001). These markers were also elevated in HPV-positive patients, who had a significantly lower mean age. HPV positivity varied significantly among pap smear groups [$\chi^2(7)=873.17$, p<0.001]; this rate increased with the severity of cytology, rising from 5.2% in benign cases to 49.9% in ASC-US, 72.9% in HSIL, and 100% in high-grade lesions.

Spearman's correlation analysis revealed moderate-to-strong positive correlations for NLR with smear results, colposcopy findings, and HPV positivity (r=0.514–0.665; all p<0.001). For SII, similar but weaker associations were found (r=0.425–0.554). PLR exhibited weak correlations (r=0.261–0.329). Age showed weak and negative correlations with NLR and SII (Table 2).

In univariate analysis, NLR, PLR, SII, and HPV positivity proved to be significant predictors of each outcome (all

**Table 1.** Demographic, clinical, and laboratory characteristics of the study population.

| Variable | n/median (IQR) | %/(min–max) | Shapiro-Wilk p-value |
|---|---|---|---|
| Demographic characteristics | | | |
| Number of patients | 2,758 | – | – |
| Age (years) | 37 (30–42) | 21–53 | <0.001 |
| Parity—nulliparous | 1,069 | 38.8 | – |
| Parity—multiparous | 1,689 | 61.2 | – |
| Hematological parameters | | | |
| NLR | 2.04 (1.61–2.53) | 0.30–15.72 | <0.001 |
| PLR | 125.9 (101.4–156.9) | 32.4–1,200.0 | <0.001 |
| SII (neutrophil × platelet/lymphocyte) | 554.1 (414.1–741.6) | 88.3–7,104.0 | <0.001 |
| Smear results | | | |
| Benign | 1,841 | 66.8 | – |
| ASC-US | 503 | 18.2 | – |
| LSIL | 312 | 11.3 | – |
| ASC-H | 40 | 1.5 | – |
| HSIL (CIN2–3) | 48 | 1.7 | – |
| AGC/AIS/SCC–adenocarcinoma | 14 | 0.5 | – |

Continuous variables are presented as median (IQR, Q1–Q3). All continuous variables showed non-normal distribution (Shapiro-Wilk p<0.001). Categorical variables are expressed as n (%). IQR: interquartile range; NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; SII: systemic immune-inflammation index (neutrophil × platelet/lymphocyte); LSIL: low-grade squamous intraepithelial lesion; ASC-H: atypical squamous cells, cannot exclude HSIL; HSIL: high-grade squamous intraepithelial lesion; CIN2–3: cervical intraepithelial neoplasia grades 2–3; AGC: atypical glandular cells; AIS: adenocarcinoma in situ; SCC: squamous cell carcinoma; ASC-US: atypical squamous cells of undetermined significance.

**Table 2.** Correlation of neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, and systemic immune-inflammation index with clinical variables.

| Marker | Age r (p) | Smear result r (p) | Colposcopy r (p) | HPV positivity r (p) |
|---|---|---|---|---|
| NLR | -0.048 (0.011)* | 0.514 (<0.001)*** | 0.572 (<0.001)*** | 0.665 (<0.001)*** |
| PLR | -0.034 (0.078) ns | 0.261 (<0.001)*** | 0.277 (<0.001)*** | 0.329 (<0.001)*** |
| SII | -0.038 (0.048)* | 0.425 (<0.001)*** | 0.477 (<0.001)*** | 0.554 (<0.001)*** |

Spearman correlation analysis. *p<0.05; ***p<0.001; ns: not significant. NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; SII: systemic immune-inflammation index.

p<0.001). In multivariate analysis, however, the significance of PLR was reduced. While NLR (OR 3.627) and HPV positivity (OR 1.755) were independent predictors of abnormal smear results, NLR (OR 2.245), SII (OR 1.439), and HPV positivity (OR 1.957) independently predicted advanced lesions (all p<0.001).

The ROC analysis indicated that NLR demonstrated the highest diagnostic performance in identifying high-grade lesions (AUC=0.871); the optimal cutoff value was >4.41 (sensitivity 66.7%, specificity 99.2%). SII also demonstrated acceptable performance (AUC=0.847), whereas PLR showed lower accuracy (AUC=0.741). Since the cutoff values were obtained from the same dataset, external validation is required (Table 3).

## DISCUSSION

This study reveals that systemic inflammatory markers show a statistically significant correlation with the severity of cervical cytopathology and HPV positivity in a large screening cohort. Our key finding demonstrates a progressive increase in these markers as cervical lesions progress from benign to high grade. Specifically, the NLR emerged as the strongest independently predictive marker for high-grade lesions (HSIL+), with a remarkable specificity of 99.2% at a cutoff value of 4.41.

The linear progression of HPV positivity—ranging from 5.2% in the benign group to 100% in high-grade lesions—highlights the critical role HPV performs in the development of cervical cancer. This tendency is consistent with the established understanding that high-risk HPV types promote a chronic inflammatory environment prone to neoplastic transformation[14]. Our results suggest that this inflammation is not limited to the local area, but appears to be systemic, as indicated by elevated hematological indices. The strong correlation between HPV positivity and NLR (r=0.665) suggests that the systemic inflammatory response is linked to viral load or the persistence of high-risk HPV subtypes. This situation is consistent with recent findings regarding the complementary role of HPV status in cytological smear reporting[15]. Additionally, in a study carried out by Tóth et al. involving 395 patients, the NLR was observed to be significantly associated with HPV DNA positivity and p16 expression; the NLR was found to be higher in HPV-positive patients. But its diagnostic performance remained moderate (AUC≈0.61)[16].

In a 2024 study by Qin and Zhang evaluating 324 participants, it was found that although NLR, PLR, and SII levels were significantly higher in patients with cervical lesions, the predictive power of these markers alone was restricted, and their AUC values were relatively low, which indicates that the clinical utility of these markers as triage biomarkers is limited. Compared with other studies, our study achieved an AUC of 0.871 for NLR, which is statistically significantly higher than previously reported values in the literature, including the AUC of 0.569 reported by Qin and Zhang[17]. This inconsistency may be related to our larger sample size (n=2,758) and the inclusion of a wider range of cytological abnormalities in the study.

Atlıhan et al. noted that although both NLR and PLR levels tended to increase with the severity of CIN, only NLR showed a statistically significant association, whereas PLR was not an independent predictor[18]. The lower specificity of the PLR may be attributed to the fact that platelet count is influenced by multiple factors. Consequently, the NLR appears to be a more reliable biomarker for cervical lesions. While the PLR was found to be meaningful in the univariate analysis, it likely became insignificant after adjustment due to multicollinearity with the NLR and SII. In contrast, NLR was a significant independent predictor in all models; this confirms its clinical value as a simpler and more reliable parameter compared to PLR or SII.

Our findings have significant clinical value. Given the NLR's high specificity (99.2%) at a cutoff >4.41, this indicator can serve as a meaningful "diagnostic verification" tool in triage procedures. In situations where resources are limited or as a supplement to primary HPV/cytology screening, the

**Table 3.** ROC analysis of neutrophil-to-lymphocyte ratio, platelet-to-lymphocyte ratio, and systemic immune-inflammation index for the diagnosis of high-grade cervical lesions.

| Marker | AUC (95%CI) | Optimal cut-off | Sensitivity (%) | Specificity (%) | PPV (%) | NPV (%) |
|---|---|---|---|---|---|---|
| NLR | 0.871 (0.827–0.916) | >4.41 | 66.7 | 99.2 | 76.4 | 98.7 |
| SII | 0.847 (0.800–0.895) | >818.65 | 74.5 | 83.1 | 14.5 | 98.8 |
| PLR | 0.741 (0.685–0.797) | >133.94 | 81.4 | 58.8 | 7.1 | 98.8 |

Optimal cutoff values were determined using the Youden index. AUC confidence intervals were calculated by the DeLong method. An AUC >0.80 was considered to indicate good discriminatory performance. PPV: positive predictive value; NPV: negative predictive value; NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; SII: systemic immune-inflammation index; CI: confidence interval; AUC: area under the curve

NLR can help determine which patients should be referred for colposcopy, potentially reducing unnecessary procedures while ensuring that high-risk patients receive prompt treatment. The availability and cost-effectiveness of complete blood counts enhance the feasibility of incorporating these markers into routine clinical practice[19].

Despite the strengths of this study, certain limitations must also be considered. First, the retrospective nature of this study complicates the identification of a definitive causal relationship between systemic inflammation and the progression of cervical lesions. The small sample sizes in rare high-risk groups (AGC, AIS, SCC/adenocarcinoma) limit the generalizability of the findings. Also, although we have determined optimal threshold values, they were derived from the same data set; therefore, an external validation study in independent cohorts is required to verify their diagnostic performance[17].

## CONCLUSION

In conclusion, markers of systemic inflammation are associated with abnormal cervical screening results, but their clinical utility varies. Among these, NLR appears to be a reliable independent predictor of abnormal cytology and high-grade lesions, while PLR lost its significance after adjusting for covariates, and SII did not demonstrate superior performance. While these findings suggest that inflammation may play a role in cervical carcinogenesis and HPV persistence, they indicate that NLR is the most practical biomarker. Given their limited discriminatory ability, however, these markers should be considered complementary rather than primary screening tools. Future multicenter prospective studies are needed to validate these findings.

## AUTHORS' CONTRIBUTIONS

**PY:** Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing. **GT:** Conceptualization, Data curation, Formal Analysis, Writing – review & editing. **GA:** Conceptualization, Formal Analysis, Investigation, Methodology, Writing – review & editing. **AU:** Conceptualization, Resources, Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Alhudhud M, Maqsood S, Hussein ME, Shaheen R, Sarhan H, Aslam S, et al. Cervical cancer screening: a comparative study of TruScreen vs. Pap smear. BMC Womens Health. 2025;25(1):198.
2. Filho AM, Laversanne M, Ferlay J, Colombet M, Piñeros M, Znaor A, et al. The GLOBOCAN 2022 cancer estimates: data sources, methods, and a snapshot of the cancer burden worldwide. Int J Cancer. 2025;156(7):1336-46.
3. İlgen O, Ejder Tekgündüz S, Yoruk N, Çelebioğlu NO, Kadıoğlu BG. Serum inflammatory biomarkers to discriminate the cervical squamous intraepithelial lesions. J Basic Clin Health Sci. 2026;10(1):97-103.
4. Kakotkin VV, Semina EV, Zadorkina TG, Agapov MA. Prevention strategies and early diagnosis of cervical cancer: current state and prospects. Diagnostics (Basel). 2023;13(4):610.
5. Stuebs FA, Schulmeyer CE, Mehlhorn G, Gass P, Kehl S, Renner SK, et al. Accuracy of colposcopy-directed biopsy in detecting early cervical neoplasia: a retrospective study. Arch Gynecol Obstet. 2019;299(2):525-32.
6. Buskwofie A, David-West G, Clare CA. A review of cervical cancer: incidence and disparities. J Natl Med Assoc. 2020;112(2):229-32
7. Mai Q, Yang X, Cheng H, Wu G, Wu Z. Prevalence and genotype distribution of human papillomavirus among women with cervical lesions in Shenzhen city, China. Hum Vaccin Immunother. 2021;17(4):965-71.
8. Sun C, Zhang J, Pan L, Yao S, Zhang F, Ji L, et al. Innovative nomogram for cervical cancer prediction: integrating high-risk HPV infection, p53 genotype, and blood routine parameters. Front Oncol. 2025;15:1541928.
9. Xu M, Wu Q, Cai L, Sun X, Xie X, Sun P. Systemic inflammatory score predicts overall survival in patients with cervical cancer. J Cancer. 2021;12(12):3671-7.
10. Ji Y, Wang H. Prognostic prediction of systemic immune-inflammation index for patients with gynecological and breast cancers: a meta-analysis. World J Surg Oncol. 2020;18(1):197.
11. Wu Y, Tu C, Shao C. Inflammatory indexes in preoperative blood routine to predict early recurrence of hepatocellular carcinoma after curative hepatectomy. BMC Surg. 2021;21(1):178.
12. Yin Y, Zhang Y, Li L, Zhang S, Liu N, Yuan S. Prognostic value of pretreatment lymphocyte-to-monocyte ratio and development of a nomogram in breast cancer patients. Front Oncol. 2021;11:650980.
13. Yıldız HA, Değer MD, Aslan G. Prognostic value of preoperative inflammation markers in non-muscle invasive bladder cancer. Int J Clin Pract. 2021;75(6):e14118.
14. Cho HY, Yim S, Yeo I, Kyung MS. The prevalence of abnormal Pap smear in women with pelvic inflammatory disease and determine the risk factors of cervical intra-epithelial neoplasia 2/3. Eur J Gynaecol Oncol. 2021;42(1):105-9.
15. Ali M, Sinha R, Kumar A, Karim S, Irfan M, Kumar S, et al. HPV DNA status and clinical history of patients are supplements

for accurate reporting of the cytological Pap smear. Sci Rep. 2024;14(1):17486.

16. Tóth Z, Lőczi L, Sebők B, Merkely P, Keszthelyi E, Lintner B, et al. Neutrophil/lymphocyte ratio (NLR) as a predictive marker for p16 positivity and cervical cancer progression: insights from the SCOPE study. Cancers (Basel). 2025;17(6):921. https://doi.org/10.3390/cancers17060921
17. Qin L, Zhang L. The predictive value of serum inflammatory markers for the severity of cervical lesions. BMC Cancer. 2024;24:780.
18. Atlıhan U, Yavuz O, Ersak B, Ata C, Bildacı TB, Erkilinc S, et al. The role of serum inflammatory markers in determining the severity of cervical lesions. J Obstet Gynaecol Res. 2025;51(3):e16278. https://doi.org/10.1111/jog.16278
19. Novin K, Fadavi P, Mortazavi N, Arefpour AM, Dehghan Manshadi HR, Shahin M, et al. Predictive value of pretreatment hematological parameters for response to definitive chemoradiotherapy in locally advanced cervical cancer: a prospective study. Int J Cancer Manag. 2024;17(1):e143466. https://doi.org/10.5812/ijcm-143466

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260753

# Artificial intelligence in acne vulgaris patients

Tuğba Atak[1]*, Tuba Betül Karadeniz[1]

## SUMMARY

**OBJECTIVE:** The aim of this study was to evaluate the experiences and confidence levels of acne patients using artificial intelligence tools. This study investigates whether patients view artificial intelligence as a standalone resource or a facilitator of physician collaboration to support the ethical implementation of healthcare technologies.

**METHODS:** This cross-sectional, descriptive study included acne patients aged ≥18 years with internet access. Alongside physician-led clinical evaluations, a structured questionnaire was administered.

**RESULTS:** Among 123 patients (81.3% female), 45.5% utilized artificial intelligence tools for general health purposes before consulting a physician, with acne-specific artificial intelligence usage being significantly driven by the perception of rapid access ($r=0.490$, $p<0.001$). Despite this interest, 67.5% insisted on professional validation for artificial intelligence advice, and notably, no significant correlation was found between clinical acne score and artificial intelligence utilization for acne management. Additionally, 55.3% of participants expressed concern about the risk of artificial intelligence making incorrect diagnoses, and 48.8% stated that they were reluctant to share personal data, such as photos, with these systems. Of those who prefer artificial intelligence, 53.7% use it for "rapid and easy access," while 52% stated that they "would never use it without consulting a doctor."

**CONCLUSION:** The integration of artificial intelligence in dermatology necessitates a physician-led, transparent approach that prioritizes digital health literacy and ensures these tools serve as supervised supplements rather than a replacement for human-centered care.

**KEYWORDS:** Artificial intelligence. Acne. Patient. Perspective.

## INTRODUCTION

Characterized by visible primary lesions and potential permanent scarring, acne is a prevalent inflammatory disease that imposes a substantial physical and psychosocial burden, profoundly impairing quality of life across both adolescent and adult populations[1].

Artificial intelligence (AI), particularly deep learning, has significantly advanced healthcare across various specialties, including dermatology, by augmenting clinical expertise rather than replacing physicians. However, the ultimate clinical success of these technologies depends not only on technical accuracy but also on patient trust and perceptions[2-5].

The majority of existing literature focuses on the algorithmic performance of AI, leaving a gap in our understanding of the attitudes and behaviors of the patients who use this technology. This study aims to directly investigate the experiences and confidence levels of acne patients using AI tools. Our findings will inform future patient education strategies by clarifying whether patients perceive AI as a simple information source or as a facilitator of physician collaboration. By documenting the patient perspective, this research provides the empirical evidence necessary to support the ethical and safe implementation of healthcare technologies.

## METHODS

This cross-sectional, descriptive study was conducted among volunteer patients diagnosed with acne at the Dermatology outpatient clinic of Kırıkkale University Faculty of Medicine between December 2025 and February 2026. Following a detailed explanation of the study objectives, informed consent was obtained from all participants. Data were collected via a purpose-designed, nongraded descriptive questionnaire, which underwent a pilot study to ensure clarity and comprehensibility. The survey instrument was developed by the investigators specifically for this study population and has not been utilized in previous publications.

Inclusion criteria required participants to be aged 18 years or older, possess adequate cognitive health to complete the survey, and have access to both the internet and a smart device (smartphone, tablet, or computer).

[1]Kırıkkale University, Faculty of Medicine, Dermatology Clinic – Kırıkkale, Türkiye.
*Corresponding author: tugba.atak@dr.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on February 07, 2026. Accepted on July 11, 2026.
Scientific Editor: José Maria Soares Júnior

The patient questionnaire comprised six main sections: socio-demographic characteristics, clinical history, AI experience (frequency and primary objectives), the perceived impact of AI on the patient–physician relationship and consultation frequency, attitudes and trust regarding AI advice (adherence, necessity of dermatologist validation, and error concerns), and perceptions of data security and AI's future role in acne management. The complete list of the survey questions is provided in Supplementary Appendix 1.

During the clinical assessment, the physician recorded the patients' acne severity using the Global Acne Grading System (GAGS)[6], acne type, the presence of scarring, and history of previous treatments.

The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Participants' identity information was kept completely anonymous, and their consent was obtained by written/verbal declaration that the data would be used solely for scientific purposes. The participants were reminded that they had the right to refuse to participate in the study or to withdraw at any stage. The study protocol was reviewed and approved by the Kırıkkale University Non-Interventional Clinical Research Ethics Committee (Decision No: 2025.12.13, Date: 10.12.2025).

### Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics software (Version 20.0). The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Continuous data were expressed as mean±standard deviation (SD), while categorical variables were presented as percentages. Categorical variables were compared using the chi-square test. For continuous variables, Student's t-test was utilized for parametric data, whereas the Mann-Whitney U and Kruskal-Wallis tests were employed for non-parametric comparisons. Pearson correlation analysis was conducted to examine the relationships between acne-related factors (e.g., AI error concern, age, and acne score). Statistical significance was set at $p<0.05$.

## RESULTS

Of the 123 patients participating in the study, 81.3% (n=100) were female and 18.7% (n=23) were male. The mean age was 23.37±5.02 years. Sociodemographic characteristics and internet usage times are shown in Table 1.

The patients' average GAGS in the clinical evaluation was 14.41±7.15. Scars were present in 76.4% of patients (Table 1). Of the methods used by patients to treat their acne, the most frequently used method was dermatologist recommendation at 74.8%. Social media recommendations ranked second at 40.7%, followed by pharmacy recommendations at 35.0%, with AI recommendations placed fourth at 8.9%. In total, 48% of patients stated that AI use would not affect their relationship with their physician.

A total of 45.5% of participants (n=56) stated that they had used AI tools for general health purposes before consulting a physician. However, when looking specifically at dermatological concerns, 60.1% of participants had never consulted AI for acne. Notably, only a small fraction of the entire cohort (4.1%, n=5) proceeded to implement treatment recommendations received directly from AI for their acne. Notably, 67.5% of participants believe that AI recommendations should always be approved by a dermatologist. Concurrently, 39.8% (n=49) of the participants stated that they would rely on AI tools as an alternative if a dermatologist were completely unreachable. Additionally, 55.3% of participants expressed concern about the risk of AI making incorrect diagnoses, and 48.8% stated that they were reluctant to share personal data, such as photos, with these systems (Figure 1). Of those who prefer AI, 53.7% use it for "rapid and easy access," while 52% stated that they "would never use it without consulting a doctor."

In the comparative analyses performed, no statistically significant difference was found between the presence of scars and AI error anxiety ($p=0.29$) or between gender and the level of trust in AI recommendations ($p=0.56$). Despite the presence of a modest positive trend between acne severity and the level of trust in AI, this relationship did not attain statistical significance ($p=0.057$, rho=0.172).

The relationships between participants' use of AI for acne and demographic, clinical, and perceptual variables were evaluated using Pearson correlation analysis (Table 2). The analysis revealed a statistically significant positive relationship between AI utilization and educational status ($r=0.192$, $p=0.034$). However, no significant correlations or associations were observed between AI usage and the duration of acne ($r=0.046$, $p=0.615$) or occupational categories ($p=0.274$). Furthermore, when evaluating lifestyle parameters, a statistically significant positive correlation was identified between AI utilization and alcohol consumption ($r=0.227$, $p=0.012$). However, no significant relationships were observed between AI usage and smoking status ($r=0.149$, $p=0.100$) or BMI ($r=-0.079$, $p=0.387$). The analysis revealed a moderate, positive, and significant relationship between AI use for acne management and AI providing rapid and easy access ($r=0.490$; $p<0.001$).

A significant correlation was identified between trust in AI recommendations and usage ($r=-0.455$; $p<0.001$). This negative correlation indicates that as the level of trust increases based on

scale scoring, the tendency to use AI also increases (1: I generally trust, 3: I don't trust them at all). Furthermore, a weak but significant negative relationship was found between age and AI usage (r=-0.200; p=0.027), indicating that younger participants preferred AI tools more frequently. A notable finding was the absence of a statistically significant relationship between the GAGS, which reflects the participants' clinical condition, and the use of AI for acne (r=-0.012; p=0.892).

## DISCUSSION

This study demonstrates that AI utilization among acne patients is independent of clinical severity (p>0.05), contrasting with literature that frequently links digital health usage to disease severity[7,8]. This independence may reveal a critical behavioral reality unique to acne patients, suggesting that the subjective motivation to engage with digital tools does not always correlate with objective clinical scores. Even individuals with clinically mild acne might experience an immediate desire for instant information, which could potentially prompt them to utilize AI interfaces. Regardless of clinical status, adoption is primarily driven by the patient's established technological approach and the demand for rapid accessibility. Consequently, patients perceive AI not as a medical necessity, but as a modern technological tool integrated into their digital routines.

Despite the proliferation of digital health platforms[9], the vast majority of acne patients in our study adhered to dermatologist recommendations. While nearly half sought AI advice prior to consultation, fewer than 5% followed it. This finding aligns with existing literature showing that 84.4% of patients

**Table 1.** Demographic, clinical characteristics, physical examination, and treatment history of acne patients.

| Variables | Category | n (%) or mean±SD |
|---|---|---|
| Age (years) [min–max] | – | 23.37±5.02 [18–44] |
| Gender | Female/male | 100 (81.3%)/23 (18.7%) |
| Body mass index (BMI, kg/m$^2$) [min–max] | – | 22.64±3.87 [22.6–39.8] |
| Education level | University and above | 98 (79.7%) |
| | High School and below | 25 (20.3%) |
| Place of residence | Urban/rural | 118 (95.9%)/5 (4.1%) |
| Smoking/alcohol | Current users (yes/occasional) | 21 (17.1%)/22 (17.9%) |
| Occupation (top 3) | Student | 76 (61.8%) |
| | Housewife | 13 (10.6%) |
| | Teacher | 10 (8.1%) |
| Internet usage | 4–8 h | 71 (57.7%) |
| | 8 h and above | 12 (9.8%) |
| Acne duration | >3 years | 48 (39.0%) |
| | 1–3 years | 27 (22.0%) |
| | <1 year | 48 (39.0%) |
| Acne type | Inflammatory | 79 (64.2%) |
| | Non-inflammatory | 37 (30.1%) |
| | Nodulocystic | 7 (5.7%) |
| Presence of scarring | Present/absent | 94 (76.4%)/29 (23.6%) |
| Acne severity score (GAGS) | Score | 14.41±7.15 |
| Classification based on GAGS | Mild–moderate | 120 (97.6%) |
| | Severe–very severe | 3 (2.4%) |
| Patient-perceived severity | Mild–moderate | 86 (70.0%) |
| | Severe (self-perception) | 37 (30.0%) |
| Treatment history | Patients with previous treatment | 83 (67.5%) |
| | Topical/systemic/combined | 43/23/19 |

SD: standard deviation; GAGS: Global Acne Grading System; BMI: body mass index.

prefer a dermatologist's diagnosis over AI—even when accuracy is equal[10] and that exclusive trust in AI applications is limited to just 1.5%[11]. Collectively, these data underscore that patient trust in physicians remains paramount.

Prior research in skin cancer screening underscores that while patients value AI for its efficiency and accuracy, they consider physician validation essential to preserving the doctor–patient relationship[12,13]. While a majority of participants highlighted the

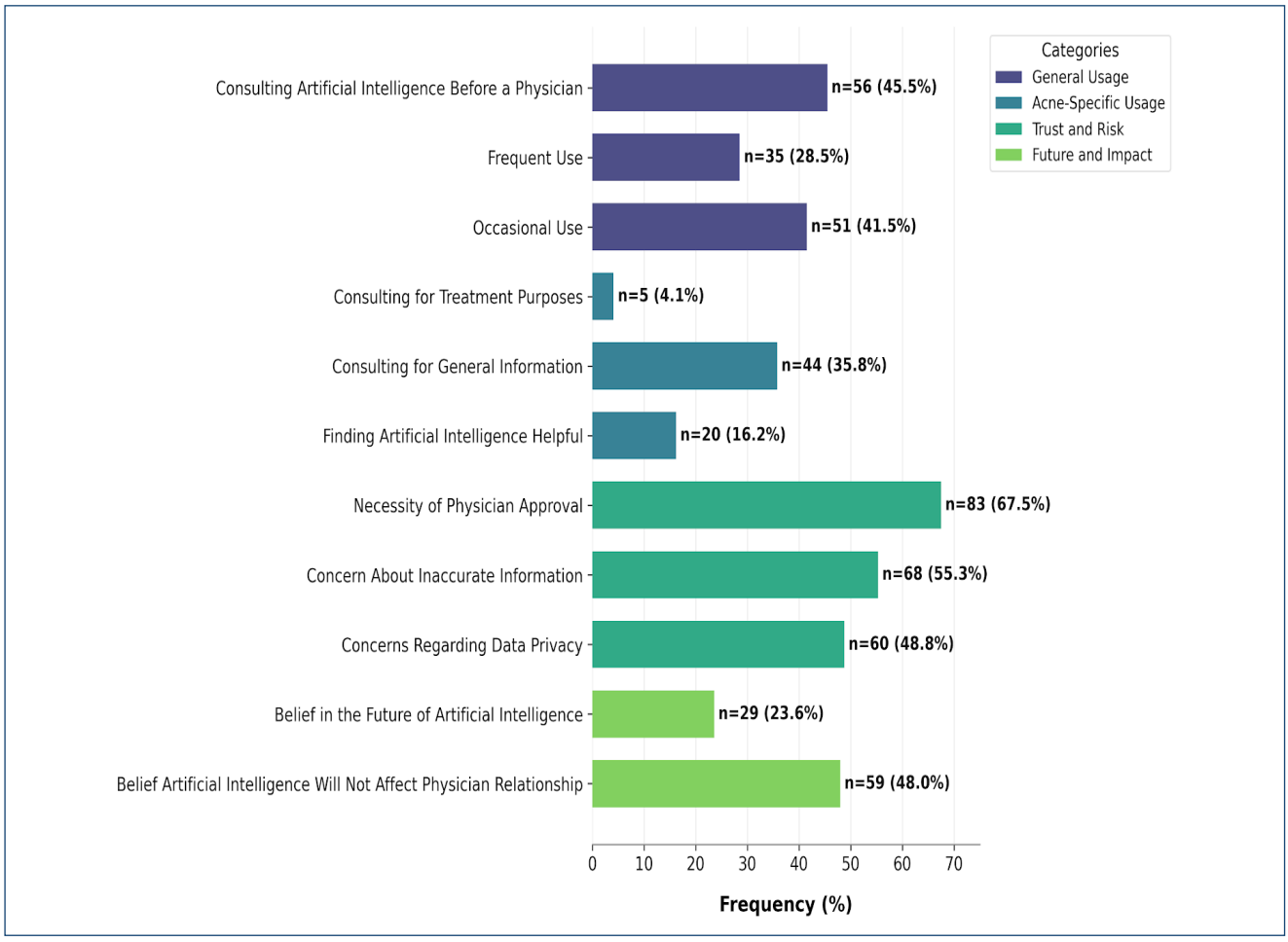



**Figure 1.** Artificial intelligence usage habits and attitudes among participants.

**Table 2.** Pearson correlation analysis between artificial intelligence use for acne and clinical and perceptual variables.

| Variables | Pearson's correlation (r) | p-value |
|---|---|---|
| Preference for speed and convenience | 0.490 | <0.001* |
| Trust in AI recommendations | -0.455 | <0.001* |
| Intention to implement AI recommendations | 0.449 | <0.001* |
| Belief in the future role of AI | 0.392 | <0.001* |
| Privacy concerns/data hesitancy | 0.229 | 0.011* |
| Age | -0.200 | 0.027* |
| AI-related concern/anxiety | -0.108 | 0.237 |
| Acne severity score (GAGS) | -0.012 | 0.892 |

*p<0.05. GAGS: Global Acne Grading System; AI: artificial intelligence.

potential advantages of AI in healthcare, including enhanced diagnostic accuracy, efficiency, time-saving benefits, and personalized care, a significant proportion of patients expressed concerns regarding data security, reliability, and the potential loss of human interaction in medical settings[14]. Mirroring these insights, 53.7% of our participants prioritized AI for its rapid accessibility, despite nearly half expressing concerns over diagnostic errors and data privacy. Crucially, the persistence of AI utilization despite these apprehensions suggests that patients often perceive immediate digital benefits, such as instant guidance, as outweighing potential data security risks.

While innovative telemedicine models can mitigate limited dermatological access[15], approximately 40% of our participants would rely on AI if a physician were unreachable. However, the "black box" nature of deep learning—the opacity of its predictive mechanisms—directly contrasts with the transparency required in standard clinical consultations[15,16]. Consequently, achieving diagnostic transparency is as vital as ensuring accuracy; AI systems must provide explainable rationales to preserve the foundational trust within the physician–patient relationship.

The findings of the present study are limited in their comparability with existing medical literature, particularly regarding acne treatment, due to a lack of data on patients' perceptions of AI. However, a more thorough examination of the available research reveals that while patients have a rudimentary understanding of the AI concept, they possess limited knowledge of its clinical applications. Demographic trends indicate that AI literacy is higher among males, younger individuals, and those pursuing higher education in science[17]. Our finding that younger age is associated with AI use aligns with the notion that younger patients are more open to technology for addressing health concerns. However, this trend carries risks, such as the potential misdirection of clinically severe acne cases toward inadequate or incorrect AI-recommended treatments, primarily due to their ease of accessibility.

Notably, a patient's familiarity with technology does not invariably correlate with their level of trust. Instead, patients emphasize the importance of AI operating under the direct supervision of a clinician, serving as a complementary tool rather than a substitute[17]. Indeed, the results of the present study indicate that 55% of respondents expressed concern regarding erroneous AI recommendations, while 67% considered clinician validation to be essential.

While literature suggests AI can enhance patient–clinician interactions by encouraging professional care[12,17,18] and reducing superfluous visits[17,19], our findings indicate that AI utilization does not actively influence the patient–physician relationship. This neutral impact aligns with previous data showing that online health information acquisition had a completely neutral effect on the patient–physician dynamic for 81.4% of individuals[20].

Limitations of our study include the limited sample size, lack of demographic diversity, and the homogeneity of the study group, which primarily consisted of young university students. It is important to note that some patients' views may be based on media perceptions or assumptions rather than actual experiences, as they lacked prior experience with AI for acne management. Furthermore, the specific AI tools or platforms utilized by the patients were not identified, precluding an analysis of how distinct algorithmic interfaces or functional capabilities might influence patient trust and satisfaction. Given that different AI modalities can uniquely influence patient trust and perception, explicitly categorizing these platforms remains a critical requirement for future research. Finally, as the study was conducted in a clinical setting, it did not account for individuals who exclusively self-medicate or do not seek professional dermatological consultation.

While further research is needed, the findings of this study have significant implications for medicine as a whole. As Clark and colleagues emphasized, AI should be used cautiously, and trends in this rapidly evolving field must be closely monitored[21]. Furthermore, while the potential benefits of AI are promising, it is imperative to address challenges such as data privacy, ethical concerns, and the necessity for ongoing collaboration between AI developers and medical professionals[22].

The outcomes of this study offer three primary insights for clinical practice: first, dermatologists should be aware that their acne patients may turn to AI tools regardless of their clinical presentation; second, it reassures dermatologists that AI acts as an initial consultation mechanism rather than a physician replacement; and third, it underscores the importance for dermatologists to guide their acne patients in digital health literacy, combining the "speed" advantage offered by AI with professional "oversight." In conclusion, the future of AI in dermatology—specifically in acne management—depends not only on technical proficiency but also on the integration of these tools into clinical workflows in a manner that is physician-controlled, transparent, explainable, and aligned with patient values.

## DECLARATION OF GENERATIVE AI AND AI-ASSISTED TECHNOLOGIES

Gemini AI Pro was used for linguistic refinement and for converting raw tabular results into the figure included in this manuscript to enhance data clarity. The authors reviewed and approved all AI-generated suggestions and take full responsibility for the content of the published article.

## AUTHORS' CONTRIBUTIONS

TA: Conceptualization, Data curation, Formal Analysis, Writing – original draft. TBK: Formal Analysis, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Harper JC. Acne vulgaris: what's new in our 40th year. J Am Acad Dermatol. 2020;82(2):526-7.
2. Hilbers D, Nekain N, Bates AT, Nunez JJ. Patients' attitudes toward artificial intelligence (AI) in cancer care: a scoping review protocol. PLoS One. 2025;20(1):e0317276.
3. Busnatu Ş, Niculescu AG, Bolocan A, Petrescu GED, Păduraru DN, Năstasă I, et al. Clinical applications of artificial intelligence-an updated overview. J Clin Med. 2022;11(8):2265.
4. Kantor J. Generative artificial intelligence in dermatology: a primer. Dermatol Clin. 2025;43(4):603-9.
5. Erul E, Aktekin Y, Danışman FB, Gümüştaş ŞA, Aktekin BS, Yekedüz E, et al. Perceptions, attitudes, and concerns on artificial intelligence applications in patients with cancer. Cancer Control. 2025;32:10732748251343245.
6. Doshi A, Zaheer A, Stiller MJ. A comparison of current acne grading systems and proposal of a novel system. Int J Dermatol. 1997;36(6):416-8.
7. Aslan Kayıran M, Karadağ AS, Alyamaç G, Çevirgen Cemil B, Demirseren DD, Aksoy H, et al. Social media use in patients with acne vulgaris: what do patients expect from social media? J Cosmet Dermatol. 2021;20(8):2556-64.
8. Ertekin SS, Salici NS, Manav Bas V, Karali MG, Ergün EZ, Avcı EB, et al. Influence of social media and internet on treatment decisions in adult female acne patients: a cross-sectional survey study. Dermatol Pract Concept. 2024;14(3):e2024156.
9. Öztekin C, Öztekin A. The association of depression, loneliness and internet addiction levels in patients with acne vulgaris. Biopsychosoc Med. 2020;14:17.
10. Wu A, Ngo M, Thomas C. Assessment of patient perceptions of artificial intelligence use in dermatology: a cross-sectional survey. Skin Res Technol. 2024;30(3):e13656.
11. McRae C, Anderson M, Turner L, Savage A, Mohammad S, Cowan R, et al. Patient perceptions of artificial intelligence integration in dermatology: a cross-sectional study of trust, comfort and equity across multiple care modalities. Skin Health Dis. 2026;6(1):35-44.
12. Nelson CA, Pérez-Chada LM, Creadore A, Li SJ, Lo K, Manjaly P, et al. Patient perspectives on the use of artificial intelligence for skin cancer screening: a qualitative study. JAMA Dermatol. 2020;156(5):501-12.
13. Lim K, Neal-Smith G, Mitchell C, Xerri J, Chuanromanee P. Perceptions of the use of artificial intelligence in the diagnosis of skin cancer: an outpatient survey. Clin Exp Dermatol. 2022;47(3):542-6.
14. Osnat B. Patient perspectives on artificial intelligence in healthcare: a global scoping review of benefits, ethical concerns, and implementation strategies. Int J Med Inform. 2025;203:106007.
15. Puri P, Comfere N, Drage LA, Shamim H, Bezalel SA, Pittelkow MR, et al. Deep learning for dermatologists: part II. Current applications. J Am Acad Dermatol. 2022;87(6):1352-60.
16. Fritsch SJ, Blankenheim A, Wahl A, Hetfeld P, Maassen O, Deffge S, et al. Attitudes and perception of artificial intelligence in healthcare: a cross-sectional survey among patients. Digital Health. 2022;8:20552076221116772.
17. Young AT, Amara D, Bhattacharya A, Wei ML. Patient and general public attitudes towards clinical artificial intelligence: a mixed methods systematic review. Lancet Digital Health. 2021;3(9):e599-611.
18. Meyer AND, Giardina TD, Spitzmueller C, Shahid U, Scott TMT, Singh H. Patient perspectives on the usefulness of an artificial intelligence-assisted symptom checker: cross-sectional survey study. J Med Internet Res. 2020;22(1):e14679.
19. Gao S, He L, Chen Y, Li D, Lai K. Public perception of artificial intelligence in medical care: content analysis of social media. J Med Internet Res. 2020;22(7):e16649.
20. Gantenbein L, Navarini AA, Maul LV, Brandt O, Mueller SM. Internet and social media use in dermatology patients: search behavior and impact on patient-physician relationship. Dermatol Ther. 2020;33(6):e14098.
21. Clark Lambert W, Grzybowski A. Dermatology and artificial intelligence. Clin Dermatol. 2024;42(3):207-9.
22. Chen M, Zhou AE, Jain N, Gronbeck C, Feng H, Grant-Kels JM. Ethics of artificial intelligence in dermatology. Clin Dermatol. 2024;42(3):313-6.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260760

# Lipedema research (2000–2025): a global bibliometric analysis with evidence of an emerging immunological paradigm

Alexandre Campos Moraes Amato[1]*, Juliana Lelis Spirandeli Amato[1], Daniel Augusto Benitti[2]

## SUMMARY

**OBJECTIVE:** The aim of this study was to characterize global scientific output on lipedema from 2000 to 2025 and identify emerging thematic trends, with emphasis on the emergence of immunological research.

**METHODS:** Publications indexed in Google Scholar, PubMed, Crossref, OpenAlex, and Scopus were retrieved using the terms "lipedema" OR "lipoedema." Records were deduplicated by normalized title matching and DOI cross-referencing. Bibliometric indicators computed included publication volume, citation metrics, field h-index, exponential growth rate, geographic distribution, author metrics, specialty classification (journal-name keyword matching supplemented by OpenAlex concept tags), and immunological content (regular-expression pattern matching against a predefined term list).

**RESULTS:** A total of 942 unique publications were identified, accumulating 13,688 citations (field h-index=60; mean 14.5 citations per article; median 2 citations per article). Annual output grew from below 12 per year through 2009 to 145 in 2025, with a doubling time of 3.6 years ($R^2$=0.90; p<0.001). The United States, Germany, and Brazil were the most productive countries. Karen Lynn Herbst (USA) was the most influential researcher (46 papers; h-index=20; 1,586 citations). Lymphology and phlebology formed the dominant classifiable domain (17.1%), with 69.3% of publications spanning multiple disciplines. Publications with immunological content totaled 30 (3.2%), grew from essentially zero before 2019 to 4.8% of annual output in 2025, and received higher mean citations than non-immunological articles (16.6 vs. 14.5).

**CONCLUSION:** Lipedema research is undergoing rapid exponential expansion driven by growing global recognition of a historically neglected condition. The accelerating immunological trend, with consistently higher citation impact, indicates an accelerating paradigm shift toward chronic adipose tissue inflammation as a unifying pathophysiological framework.

**KEYWORDS:** Lipedema. Bibliometrics. Adipose tissue. Inflammation. Biomedical research.

## INTRODUCTION

Lipedema is a chronic progressive disorder of adipose tissue distribution predominantly affecting women, with an estimated prevalence of approximately 12.3% in Brazilian women[1]. Prevalence estimates vary across populations, and no validated global estimate is currently available. First described by Allen and Hines at the Mayo Clinic in 1940[2], the condition remained largely confined to specialist lymphology literature for decades and is frequently misdiagnosed as obesity or lymphedema[3]. Over the past decade, patient advocacy, social media, and growing cross-specialty clinical recognition have driven an unprecedented expansion of the scientific literature.

Bibliometric analysis provides a rigorous quantitative framework for characterizing the intellectual structure and thematic evolution of emerging scientific fields[4]. To our knowledge, no bibliometric characterization of the lipedema literature has been published in peer-reviewed journals.

This study has two objectives: (1) to characterize global scientific output on lipedema from 2000 to 2025, including growth dynamics, leading researchers, geographic distribution, and specialty representation; and (2) to identify and contextualize the emergence of immunological research as an accelerating thematic strand with potential paradigm-shifting implications for the mechanistic understanding and treatment of lipedema.

## METHODS

### Data sources and search strategy

We conducted a systematic multi-database search in May 2026 across Google Scholar (via SerpAPI), PubMed, Crossref, OpenAlex, and Scopus. We applied the search terms "lipedema" OR "lipoedema" to title fields without date or language restrictions (no language filter was applied at any stage).

[1]Amato Duo – São Paulo (SP), Brazil.
[2]Medical Valens Center – São Paulo (SP), Brazil.
*Corresponding author: alexandre@amato.com.br
Conflicts of interest: the authors declare there are no conflicts of interest. Funding: none.
Received on May 18, 2026. Accepted on July 05, 2026.
Scientific Editor: Roseli Nomura

The search date of May 2026 allowed complete capture of all 2025 publications. Although this analysis focuses on the modern growth period (2000 to 2025), during which 96% of the corpus was published, the search captured all available records dating back to the first description of the condition in 1940 to characterize the field's historical origins and full growth trajectory. Fewer than 5% of records (41 of 942) predate 2000; these were retained for historical context, and all bibliometric indicators are reported for the full retrieved corpus unless otherwise specified.

### Deduplication and dataset construction

We combined and deduplicated records in two sequential steps: (1) normalized title matching (lowercase conversion, punctuation removed) and (2) DOI-based cross-referencing. When duplicate records were identified, we retained the record with the highest-citation count or most complete metadata; because the highest-citation record is often the oldest indexed version, this rule may modestly favor earlier records, without affecting corpus membership. The final dataset comprised 942 unique publications spanning 1940 to 2025. Supplementary metadata (institutional affiliations, open-access status, thematic concept tags) was retrieved from the OpenAlex API for 511 records with valid DOIs (54.2% of the corpus).

### Bibliometric indicators

We computed the following indicators: annual publication volume; total and mean citations per article; field h-index (the largest value h such that h articles each received at least h citations); exponential doubling time (ordinary least-squares regression on log-transformed annual publication counts from 2010 to 2024); per-author total papers, total citations, and h-index within the corpus; and country of first-author affiliation via a curated author-to-country mapping database. Author names were normalized using a curated dictionary of canonical names and their variants (e.g., "Herbst KL," "Herbst K," and "Herbst K.L." consolidated to a single identity), combining exact and case-insensitive string matching with manual curation informed by co-authorship networks and institutional affiliation; ORCID identifiers were used to confirm identity when available. Despite these steps, some residual author-merging or author-splitting error is unavoidable, particularly for common surnames.

### Specialty classification

We assigned each article to a specialty domain with a deterministic two-stage procedure implemented as a single reproducible script. Stage 1 matched the journal title (case-insensitive substring) against curated keyword sets: *lymph*, *phlebo*, *vein*/*venous*, *angiolog*, *vasa*, or *vascular* for Lymphology/Phlebology; *plast*, *aesthet*, or *reconstr* for Plastic/Aesthetic Surgery; *endocrin*, *obes*, *metabol*, *diabet*, or *nutr* for Endocrinology/Metabolism; and *dermat*, *cutan*, *skin*, or *JDDG* for Dermatology. Stage 2 resolved records with a missing or unmatched journal name using the OpenAlex primary-topic subfield. Articles matching no domain in either stage were classified as Multidisciplinary/Other. Grouping lymphology, phlebology, and vascular medicine into a single domain reflects the historical home of lipedema within angiology and lymphology, particularly in the phlebology and vascular journals that dominate the corpus.

### Immunological content identification

We identified immunological content by regular-expression pattern matching applied to article titles against a predefined term list: *inflamm*, *macrophage*, *fibrosis*, *cytokine*, *mast cell*, *immun*, *T cell*, *interleukin*, and *NF-kB*. This title-based approach was chosen for computational efficiency and cross-database consistency; however, we acknowledge that it necessarily underestimates the true proportion of immunologically relevant publications, as studies may address immunological pathways without using these specific terms in their titles. Reported immunological proportions should therefore be interpreted as a conservative lower bound and as an index of the visibility of immunological framing in the titled literature rather than of its total prevalence.

### Statistical analysis

All analyses were performed in Python 3.9.23 (pandas 2.3.3, numpy 2.0.2, matplotlib 3.9.4, networkx 3.2.1). No commercial software was used. Analysis scripts are available from the corresponding author upon reasonable request.

## RESULTS

### Publication growth

The corpus comprised 942 unique publications spanning 1940 to 2025. Total citations were 13,688 (field h-index=60; mean 14.5 citations per article; median 2 citations per article). Annual output remained below 12 publications per year through 2009, then increased progressively from 2010, with marked acceleration from 2017, reaching 145 publications in 2025 (Figure 1). Exponential regression on data from 2010 to 2024 yielded a doubling time of 3.6 years ($R^2$=0.90; p<0.001). Publications from 2020 to 2025 account for 59.3% of the entire corpus.

A first major inflection occurred around 2010, coinciding with landmark reviews by Child et al.[5] (329 citations) and Fife et al.[3] (223 citations), followed by further high-impact contributions from Forner-Cordero et al.[6] (237 citations) and Kruppa et al.[7] (187 citations). Growth accelerated further from 2017, concurrent with the publication of formal clinical practice guidelines in the United States[8] and Europe[9].

## Leading researchers

A total of 2,942 unique authors contributed to the corpus after name normalization. Karen Lynn Herbst (USA) was the most influential by all metrics: 46 publications, 1,586 citations, and an h-index of 20 within the lipedema corpus, accounting for 11.6% of all citations. Table 1 shows the top eight researchers by corpus h-index, illustrating the concentration of citation

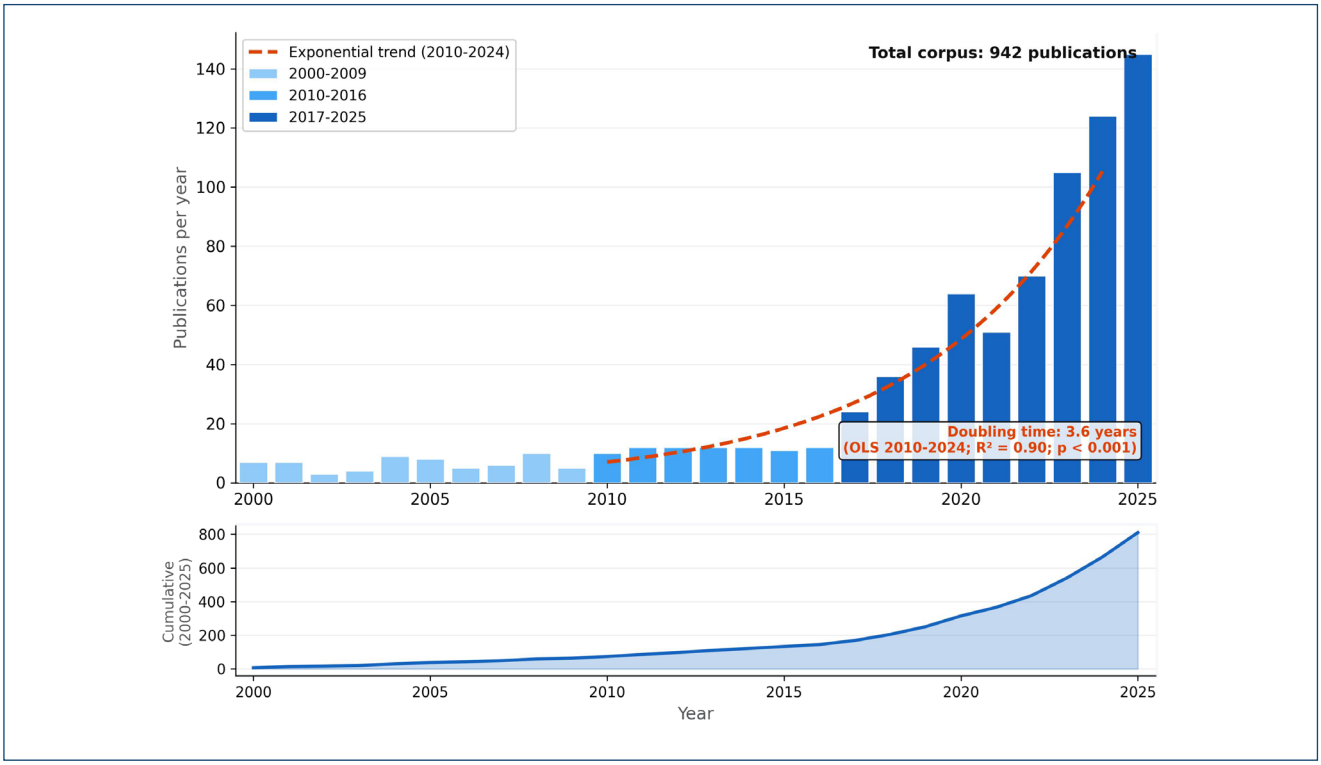



**Figure 1.** Annual publication output in lipedema research, 2000 to 2025. Bars represent the number of unique publications per year across all five databases after deduplication and are color-coded by era (2000 to 2009, 2010 to 2016, and 2017 to 2025) to highlight the successive phases of accelerating output; this era key is shown at the top of the panel. The dashed curve shows the fitted exponential trend (ordinary least-squares regression on log-transformed counts, 2010 to 2024; doubling time 3.6 years; $R^2$=0.90; p<0.001). Total corpus: 942 publications (1940 to 2025); 13,688 total citations; field h-index=60. Search conducted May 2026.

**Table 1.** Top researchers by corpus-specific h-index within the lipedema literature (1940–2025). h-index values are computed within the lipedema corpus only and do not reflect career-level h-indices.

| Rank | Author | Country | Papers | Total citations | h-index* |
|---|---|---|---|---|---|
| 1 | Karen Lynn Herbst | USA | 46 | 1,586 | 20 |
| 2 | Uwe Wollina | Germany | 21 | 644 | 12 |
| 3 | Győző Szolnoky | Hungary | 23 | 1,093 | 10 |
| 4 | Rachael L. Crescenzi | USA | 19 | 486 | 10 |
| 5 | Alexandre C. M. Amato | Brazil | 28 | 279 | 9 |
| 6 | Lajos Kemény | Hungary | 13 | 639 | 8 |
| 7 | Wolfgang Schmeller | Germany | 9 | 566 | 8 |
| 8 | Stefan Rapprich | Germany | 9 | 448 | 8 |

*Corpus-specific h-index (see table title). Ties in h-index are broken by total citations within the corpus. Career-level h-indices are not reported, as they are not uniformly available across all databases for all authors and would introduce an inconsistent metric. Data as of May 2026.

impact relative to publication volume. Paula Michelle Donahue (USA), who led a landmark national survey of patient symptoms in 707 women[10], ranks among the top contributors by h-index. Alexandre Campos Moraes Amato (Brazil), with 28 publications, is the most productive Brazilian contributor. Mean authorship per article was 3.6 (median 3); single-author papers comprised 27.0% of the corpus.

### Geographic distribution

The United States (58 publications), Germany (56), and Brazil (54) were the most productive countries, followed by Italy (31), Hungary (21), and Spain (12). An observational study of 1,044 Italian patients with lipedema further illustrates the geographic breadth of clinical research activity[11]. Citation cohesion (the proportion of a country's articles that cite at least one domestic article within the corpus, computed using the same first-author attribution as the productivity counts) was highest for Germany (88%), followed by Italy (79%), the United States (74%), and Brazil (74%). These estimates rest on small numbers of analyzable articles per country (14 to 38) and should be interpreted cautiously. We initially hypothesized that a lower cohesion in Brazil might reflect temporal lag, as the mean publication year of Brazilian articles was 2022 versus 2018 for Germany; however, under consistent attribution, Brazil does not stand out as uniquely low, and we found no reliable association between a country's mean publication year and its cohesion. Temporal lag, therefore, remains an unconfirmed hypothesis, and the observed differences are more likely multifactorial, reflecting the language of publication, journal accessibility and indexing, the maturity and size of each national literature, and patterns of international versus domestic collaboration.

### Specialty landscape

Of the 942 publications, 69.3% were classified as Multidisciplinary/Other. Lymphology and phlebology together formed the dominant classifiable domain (17.1% of all publications), followed by Endocrinology/Metabolism (5.2%), Plastic/Aesthetic Surgery (4.5%), and Dermatology (3.9%). Mean citation counts were highest for Dermatology-classified articles (47.6 per article), suggesting that dermatological contributions address topics of broad cross-specialty interest. To characterize the large Multidisciplinary/Other category, we re-queried OpenAlex concept and topic tags for its articles carrying a DOI (320 of 335 matched). At the top level, 88.1% were tagged "Medicine," with the remainder distributed across Biology (7.2%) and Psychology (3.8%); the most frequent finer-grained concepts were Internal medicine (23.8%), Surgery (22.5%), Physical therapy (16.9%), Pathology (12.8%), Endocrinology (11.2%), Dermatology (10.0%), Radiology (6.9%), and Gynecology (6.9%). This distribution confirms that the category comprises genuinely cross-specialty clinical and translational research rather than off-topic publications, reinforcing our interpretation of specialty fragmentation as a structural feature of the field.

### Emerging immunological trend

Publications with immunological content totaled 30 (3.2% of the corpus); this figure represents all immunological articles identified in the corpus, of which only key anchor studies are annotated in Figure 2. This proportion was essentially zero before 2019 (under 1 article per year on average) and grew to 4.8% of annual output in 2025 (Figure 2). Immunological articles received higher mean citations than non-immunological articles (16.6 vs. 14.5), despite being predominantly recent publications. The most cited immunological publication, Al-Ghadban et al. (169 citations)[12], documented markedly elevated macrophage density, dilated lymphatic microvessels, and angiogenesis in lipedema adipose tissue, providing the first direct histological evidence of an immune component in the disease.

## DISCUSSION

### Exponential growth of a neglected field

This 3.6-year doubling time is roughly two to three times faster than the growth of the overall biomedical literature, which has historically doubled about every 9 years[13], and it reflects a genuine epistemic transition from an isolated specialist topic to a mainstream clinical concern. This benchmark is derived from data through the early 2010s, and contemporary biomedical growth may differ; nonetheless, the gap is large enough that the qualitative conclusion is robust. The 2021 publication of the Standards of Care for lipedema in the United States[8] and the 2024 S2k guideline in Germany[9] were both products and accelerants of this process. Remarkably, 59.3% of all publications in the field's 85-year history appeared in just the most recent six years (2020 to 2025), underscoring the recency and speed of this shift. This concentration has a practical corollary: the clinical guidelines, therapeutic trials, and mechanistic studies that now shape practice were nearly all produced within a single recent window, so the evidence base is simultaneously expanding and immature, and current recommendations rest on a literature that is still consolidating. The composition of authorship reflects this transitional stage: single-author papers still comprised 27.0% of the corpus alongside a mean of 3.6 authors per article, consistent with a field moving from individual clinical reports toward multidisciplinary team science.

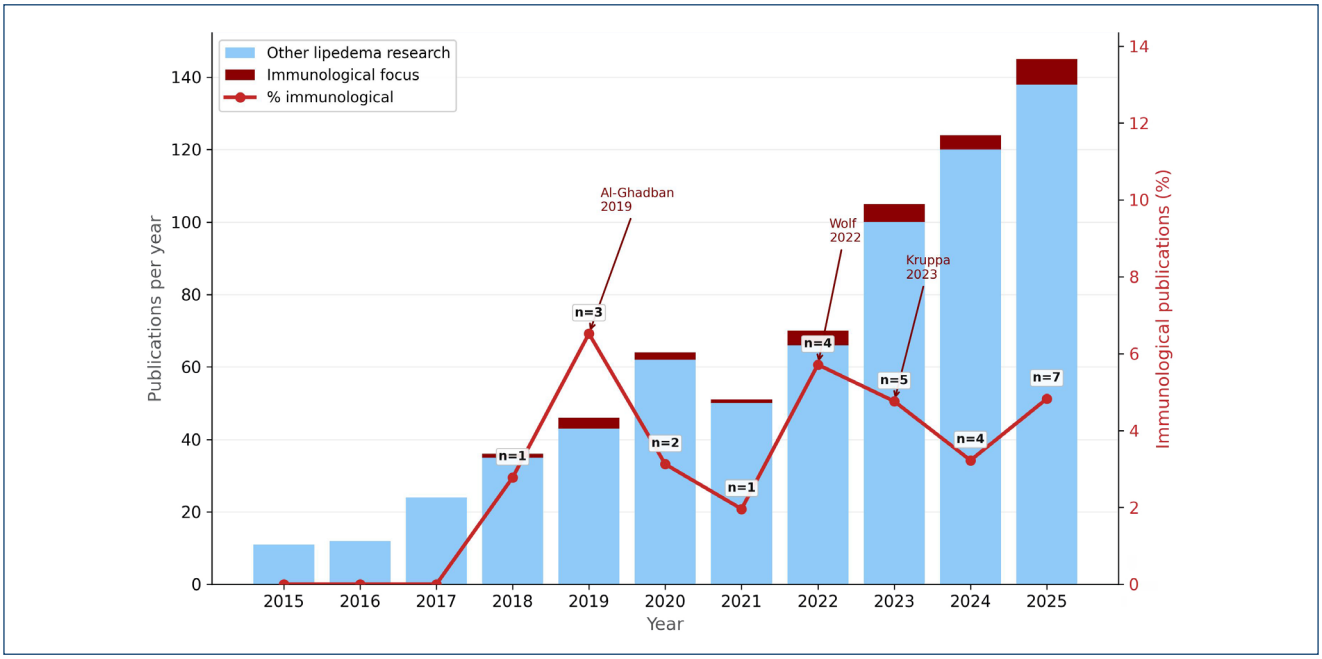



**Figure 2.** Emerging immunological trend in lipedema research, 2015 to 2025. Stacked bars represent annual publication volume: light blue, non-immunological publications; dark red, publications with immunological content (identified by regular-expression pattern matching of article titles against a predefined term list including *inflamm*, *macrophage*, *fibrosis*, *cytokine*, *mast cell*, *immun*, *T cell*, *interleukin*, and *NF-kB*). The red line (right axis) shows the percentage of immunological content per year; the label at each point gives the absolute number of immunological publications in that year (from n=1 in 2018 to n=7 in 2025), which is also encoded by the height of the dark-red bar segments. Only the key publications that anchored the immunological turn are annotated (Al-Ghadban et al.[12]; Wolf et al.[18]; and Kruppa et al.[17]); these annotations mark landmark studies and are not an exhaustive list of the 30 immunological articles. Total immunological publications: 30 (3.2% of corpus); 4.8% of 2025 output.

## Geographic leadership

The United States' dominance reflects the foundational Mayo Clinic contributions of Allen, Hines, and Wold[2,14] and the singular research program of Karen Lynn Herbst, whose 11.6% share of total citations is extraordinary in a corpus of 942 publications. Germany's leading position reflects its lymphology infrastructure and early adoption of liposuction as a therapeutic modality[15]. Brazil's emergence as the third most productive country, largely since 2018, demonstrates rapid growth outside traditional centers, supported by national prevalence data showing approximately 12.3% prevalence in Brazilian women[1] and a formal consensus statement on diagnosis and treatment published in 2025[16].

## Specialty fragmentation

The 69.3% multidisciplinary classification rate is a structural feature of the field, not a methodological artifact. Patients with lipedema are encountered by lymphologists, phlebologists, endocrinologists, plastic surgeons, dermatologists, and general practitioners, and the literature reflects this fragmentation. Among classifiable publications, lymphology and phlebology hold the largest share (17.1%) and the longest historical claim, encompassing the original diagnostic framework and the mainstay conservative treatment. The growth of collaborative authorship groups in the contemporary period suggests a shift toward multidisciplinary team-based management.

## Immunology as an emerging paradigm

The acceleration of immunological publications, combined with higher citation impact despite shorter accrual time, is consistent with paradigm formation. Al-Ghadban et al.[12] provided the first direct tissue-level evidence of an immune component, documenting macrophage infiltration and dilated lymphatic microvessels in lipedema adipose tissue. Subsequent histological studies have confirmed that adipocyte hypertrophy is accompanied by progressive subcutaneous fibrosis[17] and that a distinct M2 (alternatively activated) macrophage infiltrate decisively influences adipocyte differentiation in lipedema[18]. Elevated vascular endothelial growth factor C (VEGF-C) levels and macrophage infiltration have been documented even in cases without overt clinical lymphedema[19], suggesting that immune activation precedes rather than follows lymphatic failure. Comparative molecular analyses

further reveal distinct transcriptomic and histological signatures differentiating lipedema from lipohypertrophy and secondary lymphedema[20]. Lymphatic imaging in early stage disease has demonstrated functional abnormalities preceding structural changes[21], consistent with the hypothesis that perilymphatic immune infiltration drives lymphatic dysfunction as a downstream consequence. The hormonal dimension of this immune activation is supported by molecular studies showing estrogen-dependent upregulation of adipogenic and inflammatory pathways in lipedema adipose-derived stem cells[22]. Collectively, these converging findings support chronic low-grade adipose tissue inflammation, driven by macrophage infiltration and immune–lymphatic interaction, as the unifying pathophysiological framework that may connect the specialty-specific manifestations of lipedema and define the therapeutic frontier for the coming decade[23,24].

### Limitations

Title-based deduplication may occasionally fail to merge distinct-titled versions of the same publication or to separate articles with highly similar titles. Country attribution relied on a curated mapping that covered a subset of the corpus; countries with lower mapping coverage may be underrepresented. Specialty classification by journal-name keyword is an approximation and does not capture multi-specialty journals accurately. Because our thematic search interrogated article titles, immunological content identification may miss studies addressing immunological pathways only in the abstract or full text; reported immunological proportions are therefore a conservative lower bound. Author-name normalization, although curated, cannot fully eliminate residual merging or splitting of identities, particularly for common surnames. Citation counts represent a snapshot from May 2026 and will continue to accrue; because publications from 2023 to 2025 have had substantially less time to accumulate citations, citation-based metrics are subject to citation-lag bias that conservatively biases downward the ranking of the most recent authors and articles, and reinforces that the emerging immunological signal, already showing higher mean citations despite shorter accrual time, is likely underestimated rather than inflated. Although five databases were interrogated, no single source is exhaustive (for example, Google Scholar does not index every Scopus-listed journal), so some coverage bias may remain. First author (ACMA) and one co-author (DAB) appear in the author rankings; their data are reported using the same methodology applied to all researchers, but readers should be aware of this potential bias.

## CONCLUSION

This bibliometric analysis of 942 lipedema publications (1940 to 2025) documents rapid exponential growth (doubling time 3.6 years; $R^2$=0.90), with the United States, Germany, and Brazil as the leading contributors and Karen Lynn Herbst (46 publications; h-index 20; 1,586 citations) as the most influential researcher by all metrics. More than half the entire lipedema literature was published in the most recent 6 years (2020 to 2025), confirming the field's current momentum. The specialty landscape is inherently multidisciplinary, with lymphology as the dominant classifiable tradition. The most notable finding is the accelerating emergence of immunological research, growing from essentially zero before 2019 to 4.8% of annual output in 2025, with consistently higher citation impact. Chronic adipose tissue inflammation, mediated by macrophage infiltration and immune–lymphatic interaction, may represent the unifying pathophysiological framework for the next generation of lipedema research.

## ETHICS STATEMENT

This study involved analysis of publicly available bibliographic data without collection of personally identifiable information from individual human subjects and was therefore exempt from institutional ethics committee review.

## ACKNOWLEDGMENTS

The authors used Claude (Anthropic Inc., San Francisco, CA, USA) as a writing assistance tool during manuscript preparation. All intellectual content, data analysis, interpretation of results, and final editorial decisions were made exclusively by the authors.

## AUTHORS' CONTRIBUTIONS

**ACMA:** Conceptualization, Data curation, Formal Analysis, Methodology, Software, Writing – original draft, Writing – review & editing. **JLSA:** Validation, Writing – review & editing. **DAB:** Supervision, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Amato ACM, Amato FCM, Amato JLS, Benitti DA. Lipedema prevalence and risk factors in Brazil. J Vasc Bras. 2022;21:e20210198. https://doi.org/10.1590/1677-5449.202101981
2. Allen EV, Hines EA Jr. Lipedema of the legs: a syndrome characterized by fat legs and orthostatic edema. Proc Staff Meet Mayo Clin. 1940;15:184-7.
3. Fife CE, Maus EA, Carter MJ. Lipedema: a frequently misdiagnosed and misunderstood fatty deposition syndrome. Adv Skin Wound Care. 2010;23(2):81-92. https://doi.org/10.1097/01.ASW.0000363503.92360.91
4. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. J Bus Res. 2021;133:285-96. https://doi.org/10.1016/j.jbusres.2021.04.070
5. Child AH, Gordon KD, Sharpe P, Brice G, Ostergaard P, Jeffery S, et al. Lipedema: an inherited condition. Am J Med Genet A. 2010;152A(4):970-6. https://doi.org/10.1002/ajmg.a.33313
6. Forner-Cordero I, Szolnoky G, Forner-Cordero A, Kemény L. Lipedema: an overview of its clinical manifestations, diagnosis and treatment of the disproportional fatty deposition syndrome – systematic review. Clin Obes. 2012;2(3-4):86-95. https://doi.org/10.1111/j.1758-8111.2012.00045.x
7. Kruppa P, Georgiou I, Biermann N, Prantl L, Klein-Weigel P, Ghods M. Lipedema: pathogenesis, diagnosis, and treatment options. Dtsch Arztebl Int. 2020;117(22-3):396-403. https://doi.org/10.3238/arztebl.2020.0396
8. Herbst KL, Kahn LA, Iker E, Ehrlich C, Wright T, McHutchison L, et al. Standard of care for lipedema in the United States. Phlebology. 2021;36(10):779-96. https://doi.org/10.1177/02683555211015887
9. Faerber G, Cornely M, Daubert C, Beckmann E, Bertsch T, Breu FX, et al. S2k guideline lipedema. J Dtsch Dermatol Ges. 2024;22(8):1141-53. https://doi.org/10.1111/ddg.15513
10. Aday WA, Donahue PMC, Garza M, Sprigle A, Herrick MG, Herbst KL, et al. National survey of patient symptoms and therapies among 707 women with a lipedema diagnosis in the United States. Vasc Med. 2023;28(5):457-61. https://doi.org/10.1177/1358863X231202769
11. Patton L, Ricolfi L, Bortolon M, Gabriele G, Zolesio P, Cione E, et al. Observational study on a large Italian population with lipedema: biochemical and clinical features. Int J Mol Sci. 2024;25(3):1599. https://doi.org/10.3390/ijms25031599
12. Al-Ghadban S, Cromer W, Allen M, Ussery C, Badowski M, Harris D, et al. Dilated blood and lymphatic microvessels, angiogenesis, increased macrophages, and adipocyte hypertrophy in lipedema thigh skin and fat tissue. J Obes. 2019;2019:8747461. https://doi.org/10.1155/2019/8747461
13. Bornmann L, Mutz R. Growth rates of modern science: a bibliometric analysis based on the number of publications and cited references. J Assoc Inf Sci Technol. 2015;66(11):2215-22. https://doi.org/10.1002/asi.23329
14. Wold LE, Hines EA Jr, Allen EV. Lipedema of the legs: a syndrome characterized by fat legs and edema. Ann Intern Med. 1951;34(5):1243-50. https://doi.org/10.7326/0003-4819-34-5-1243
15. Rapprich S, Dingler A, Podda M. Liposuction is an effective treatment for lipedema – results of a study with 25 patients. J Dtsch Dermatol Ges. 2011;9(1):33-40. https://doi.org/10.1111/j.1610-0387.2010.07504.x
16. Amato ACM, Peclat APRM, Kikuchi R, Souza AC, Silva MTB, Oliveira RHP, et al. Brazilian consensus statement on lipedema using the Delphi methodology. J Vasc Bras. 2025;24:e20230183. https://doi.org/10.1590/1677-5449.202301832
17. Kruppa P, Gohlke S, Lapinski K, Garcia-Carrizo F, Soultoukis AG, Infanger M, et al. Lipedema stage affects adipocyte hypertrophy, subcutaneous adipose tissue inflammation and interstitial fibrosis. Front Immunol. 2023;14:1223264. https://doi.org/10.3389/fimmu.2023.1223264
18. Wolf S, Rannikko JH, Virtakoivu R, Cinelli P, Felmerer G, Burger A, et al. A distinct M2 macrophage infiltrate and transcriptomic profile decisively influence adipocyte differentiation in lipedema. Front Immunol. 2022;13:1004609. https://doi.org/10.3389/fimmu.2022.1004609
19. Felmerer G, Stylianaki A, Hollmen M, Strobel P, Stepniewski A, Wang A, et al. Increased levels of VEGF-C and macrophage infiltration in lipedema patients without corresponding lymphangiogenesis and clinical lymphedema. Sci Rep. 2020;10:10947. https://doi.org/10.1038/s41598-020-67987-3
20. Atzigen J, Burger A, Grunherz L, Barbon C, Felmerer G, Giovanoli P, et al. A comparative analysis to dissect the histological and molecular differences among lipedema, lipohypertrophy and secondary lymphedema. Int J Mol Sci. 2023;24(8):7591. https://doi.org/10.3390/ijms24087591
21. Rasmussen JC, Aldrich MB, Fife CE, Herbst KL, Sevick-Muraca EM. Lymphatic function and anatomy in early stages of lipedema. Obesity (Silver Spring). 2022;30(7):1391-400. https://doi.org/10.1002/oby.23458
22. Al-Ghadban S, Isern SU, Herbst KL, Bunnell BA. The expression of adipogenic marker is significantly increased in estrogen-treated lipedema adipocytes differentiated from adipose-derived stem cells. Biomedicines. 2024;12(5):1042. https://doi.org/10.3390/biomedicines12051042
23. Amato ACM. The evolutionary theory of lipedema: a perspective on energy storage and chronic inflammation. Cureus. 2025;17(7):e88809. https://doi.org/10.7759/cureus.88809
24. Felmerer G, Stylianaki A, Haegerling R, Wang A, Strobel P, Hollmen M, et al. Adipose tissue hypertrophy, an aberrant biochemical profile and distinct gene expression in lipedema. J Surg Res. 2020;253:294-303. https://doi.org/10.1016/j.jss.2020.03.055

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260783

# Night eating syndrome in women with gestational diabetes mellitus: a comparative cross-sectional study

Esra Keles[1]*, Leyla Kaya[2], Cansu Ergenç Özdaş[3], Zahide Kaya[4]

## SUMMARY

**OBJECTIVE:** The aim of this study was to determine the prevalence of night eating syndrome among women with gestational diabetes mellitus compared with healthy pregnant controls, and to evaluate its associations with sleep quality, healthy lifestyle behaviors, fatigue severity, physical activity, perceived social support, and neonatal outcomes.

**METHODS:** This single-center, comparative cross-sectional study enrolled 615 pregnant women at a tertiary hospital. Night eating syndrome was identified using the Night Eating Questionnaire ≥25. Sleep quality, lifestyle behaviors, fatigue, physical activity, and perceived social support were assessed with the Pittsburgh Sleep Quality Index, Healthy Lifestyle Behaviors Scale, Fatigue Severity Scale, International Physical Activity Questionnaire, and Multidimensional Scale of Perceived Social Support. Group comparisons, and multivariable logistic regression with Benjamini-Hochberg correction were performed.

**RESULTS:** Night eating syndrome prevalence was higher in women with gestational diabetes mellitus than in controls (p<0.001). The gestational diabetes mellitus group exhibited worse sleep (Pittsburgh Sleep Quality Index 17.2±4.0 vs. 8.3±5.7), lower Healthy Lifestyle Behaviors Scale scores, greater fatigue, and reduced social support (all p<0.001). Higher Healthy Lifestyle Behaviors Scale and physical activity were protective against night eating syndrome. Night eating syndrome was independently associated with a nearly fourfold increase in neonatal intensive care unit admission (adjusted odds ratio 3.82; 95%CI 1.68–8.71; p=0.001) after adjustment for gestational diabetes mellitus, body mass index, sleep, fatigue, lifestyle, gestational age at birth, and maternal age.

**CONCLUSION:** Night eating syndrome is highly prevalent among women with gestational diabetes mellitus, clustering with poor sleep, severe fatigue, sedentariness, and reduced social support, and independently associated with adverse neonatal outcomes.

**KEYWORDS:** Diabetes. Gestational. Feeding and eating disorders. Sleep quality. Fatigue. Night eating syndrome.

## INTRODUCTION

Gestational diabetes mellitus (GDM) is among the most common metabolic complications of pregnancy, with a global prevalence ranging from 5 to 20% depending on diagnostic thresholds and population characteristics[1]. Its clinical burden extends beyond maternal hyperglycemia and encompasses preeclampsia, macrosomia, neonatal hypoglycemia, birth trauma, and increased neonatal intensive care unit (NICU) admission[2,3]. Although current guidelines emphasize medical nutrition therapy, structured physical activity, glucose self-monitoring, and pharmacotherapy when indicated, the behavioral and psychosocial dimensions of GDM remain comparatively underexplored[4].

Night eating syndrome (NES) is a circadian-related eating disorder characterized by evening hyperphagia, nocturnal ingestion, morning anorexia, and concurrent sleep disturbance[5]. NES is classified within the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5) category of "Other Specified Feeding or Eating Disorders," and the Night Eating Questionnaire (NEQ) is the most widely validated screening instrument[6]. While general-population prevalence is 1.5–5%, rates rise to 6–16% in obesity and exceed 25% among bariatric surgery candidates[7,8]. Proposed mechanisms include circadian misalignment, hypothalamic–pituitary–adrenal axis dysregulation, and altered melatonin–leptin–ghrelin signaling—pathways that plausibly intersect with the metabolic disturbances of GDM[9,10].

Despite this biological plausibility, NES in GDM has received limited systematic attention. Existing research on disordered eating in pregnancy has focused predominantly on binge eating and bulimia nervosa[11,12], leaving the prevalence, behavioral

[1]University of Health Sciences, Kartal Dr. Lütfi Kırdar City Hospital, Department of Gynecologic Oncology – Istanbul, Turkey.
[2]University of Health Sciences, Faculty of Health Sciences, Department of Midwifery – Istanbul, Turkey.
[3]Ankara Yildirim Beyazit University, Business School – Ankara, Turkey.
[4]Uskudar State Hospital, Internal Medicine Clinic – Istanbul, Turkey.
*Corresponding author: dresrakeles@hotmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 19, 2026. Accepted on July 24, 2026.
Scientific Editor: José Maria Soares Júnior

clustering, and neonatal implications of NES in GDM largely undefined. Furthermore, whether NES contributes independently to neonatal morbidity—beyond conventional GDM-related risk pathways has not been adequately examined[13,14].

Accordingly, this study aimed to (i) compare NES prevalence between women with GDM and healthy pregnant controls; (ii) characterize the psychobehavioral and clinical profile associated with NES; and (iii) identify independent predictors of NES.

## METHODS

### Study design

A single-center, comparative cross-sectional study was conducted at a tertiary obstetric referral center. The investigation conformed to the Declaration of Helsinki and was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement.

### Participants

Women diagnosed with GDM by the International Association of Diabetes and Pregnancy Study Groups criteria and healthy pregnant controls were consecutively enrolled during routine antenatal visits.

### Inclusion and exclusion criteria

Inclusion criteria were singleton pregnancy, maternal age ≥18 years, gestational age ≥24 weeks, and ability to complete self-report instruments in the study language. Exclusion criteria included pre-existing type 1 or type 2 diabetes mellitus, multiple gestations, psychiatric disorders requiring pharmacotherapy, and major fetal anomalies. The final sample comprised 615 women (271 GDM; 344 healthy controls).

### Measurements and scales

Sociodemographic, obstetric, and anthropometric data (maternal age, gravidity, parity, planning status, gestational age at assessment, pre-pregnancy and current weight, body mass index [BMI], gestational weight gain) were retrieved from medical records and verified by structured interview.

NES was assessed using the 14-item NEQ (range 0–52); a score ≥25 defined clinically significant NES[15]. Sleep quality was measured using the Pittsburgh Sleep Quality Index (PSQI; range 0–21), with scores>5 indicating poor sleep[16]. Fatigue was evaluated by the Fatigue Severity Scale (FSS; range 1–7), with a mean score ≥4 defining severe fatigue[17]. Health-promoting behaviors were measured by the Healthy Lifestyle Behaviors Scale (HLBS; range 29–145)[18]. Perceived social support was assessed by the Multidimensional Scale of Perceived Social Support (MSPSS; range 12–60)[19], and habitual physical activity by the International Physical Activity Questionnaire (IPAQ)[20]. Gestational age at birth, birth weight, 1- and 5-min Apgar scores, NICU admission, and respiratory support were extracted from delivery and neonatal records.

### Statistical analysis

Distributions were tested by the Shapiro-Wilk test. Continuous variables were expressed as mean±standard deviation and compared using the Mann-Whitney U test owing to non-normal distributions; categorical variables were compared using the chi-square test. Effect sizes were reported as Cohen's d for continuous variables and Cramér's V (φ) for categorical variables. Two multivariable binary logistic regression models were constructed: Model A predicting NES (NEQ≥25) and Model B predicting NICU admission. Covariates were selected a priori based on clinical relevance and bivariate associations. Model performance was evaluated by the Hosmer-Lemeshow goodness-of-fit test, Nagelkerke pseudo-$R^2$, and classification accuracy. The Benjamini-Hochberg false discovery rate correction was applied across the behavioral cluster ($q<0.05$). Two-sided $p<0.05$ was considered statistically significant. Analyses were performed using Statistical Package for the Social Sciences v26.0 (IBM Corp., Armonk, NY, USA).

## RESULTS

### Baseline characteristics

Women with GDM were significantly older, of higher gravidity, and had greater pre-pregnancy and current pregnancy weight than controls (all $p<0.001$; Table 1). They were also assessed earlier in gestation (38.1±1.9 vs. 38.9±1.7 weeks; $p<0.001$). Gestational age at birth was lower in the GDM group (38.1±1.9 vs. 39.9±1.8 weeks; $p<0.001$; d=0.97), birth weight was higher ($p=0.015$), and 1- and 5-min Apgar scores were lower ($p<0.001$). NICU admission was more frequent in the GDM cohort (18.8 vs. 10.2%; $p=0.003$; φ=0.12).

### Prevalence of night eating syndrome and group differences in psychobehavioral profile

NES (NEQ≥25) was identified in 83/615 (13.5%) of the overall cohort. Among women with GDM, 78/271 (28.8%) met NES criteria compared with 5/344 (1.5%) of controls ($p<0.001$; φ=0.49). Women with GDM demonstrated significantly poorer sleep quality (PSQI 17.2±4.0 vs. 8.3±5.7), lower HLBS scores (105.7±7.2 vs. 128.1±3.7), greater fatigue (FSS 4.26±1.32 vs. 2.28±1.11), and reduced social support (all $p<0.001$).

## Psychobehavioral and neonatal profiles by night eating syndrome status

NES-positive women did not differ from NES-negative women in maternal age but had higher pre-pregnancy weight (76.5±15.1 vs. 70.2±13.4 kg; p<0.001) and BMI (28.9±5.9 vs. 27.3±5.5 kg/m$^2$; p=0.018; Table 2). PSQI scores were markedly higher among NES-positive women (18.5±4.1 vs. 11.8±6.3; d=1.16), and all NES-positive participants exceeded the poor-sleep threshold (PSQI>5: 100 vs. 75.8%; p<0.001). The largest behavioral effect size was observed for fatigue (FSS 4.92±1.18

**Table 1.** Baseline sociodemographic, obstetric, anthropometric, and neonatal characteristics of women with gestational diabetes mellitus versus healthy pregnant controls (n=615).

| Characteristic | GDM (n=271) | Controls (n=344) | p-value† |
|---|---|---|---|
| Maternal age (years) | 31.5±5.8 | 28.2±4.9 | <0.001 |
| Gravidity | 2.8±1.6 | 2.1±1.3 | <0.001 |
| Gestational age at assessment (weeks) | 38.1±1.9 | 38.9±1.7 | <0.001 |
| Pre-pregnancy weight (kg) | 72.0±14.1 | 64.1±11.4 | <0.001 |
| Current pregnancy weight (kg) | 83.8±13.6 | 78.8±11.4 | <0.001 |
| PSQI total score | 17.2±4.0 | 8.3±5.7 | <0.001 |
| HLBS total score | 105.7±7.2 | 128.1±3.7 | <0.001 |
| FSS mean score | 4.26±1.32 | 2.28±1.11 | <0.001 |
| Gestational age at birth (weeks) | 38.1±1.9 | 39.9±1.8 | <0.001 |
| Birth weight (g) | 3325±538 | 3258±490 | 0.015 |
| Apgar 1-min | 7.28±1.03 | 7.60±0.74 | <0.001 |
| Apgar 5-min | 8.68±0.62 | 8.83±0.49 | <0.001 |
| NICU admission, n (%) | 51 (18.8) | 35 (10.2) | 0.003 |

†Mann-Whitney U for continuous; $\chi^2$ for categorical. Data are mean±SD unless otherwise indicated. FSS: Fatigue Severity Scale; GDM: gestational diabetes mellitus; HLBS: Healthy Lifestyle Behaviors Scale; NICU: neonatal intensive care unit; PSQI: Pittsburgh Sleep Quality Index; SD: standard deviation.

**Table 2.** Behavioral, psychosocial, and neonatal profiles stratified by night eating syndrome status (n=615).

| Variable | NES+ (NEQ≥25) (n=83) | NES- (NEQ<25) (n=532) | p-value† |
|---|---|---|---|
| GDM, n (%) | 78 (94.0) | 193 (36.3) | <0.001 |
| Maternal age (years) | 31.2±6.4 | 31.6±5.5 | 0.527 |
| Pre-pregnancy weight (kg) | 76.5±15.1 | 70.2±13.4 | <0.001 |
| Pre-pregnancy BMI (kg/m$^2$) | 28.9±5.9 | 27.3±5.5 | 0.018 |
| Gestational weight gain (kg) | 13.8±6.5 | 13.3±6.0 | 0.420 |
| NEQ total (0–52) | 29.1±4.2 | 12.4±6.9 | <0.001 |
| PSQI total (0–21) | 18.5±4.1 | 11.8±6.3 | <0.001 |
| Poor sleep (PSQI>5), n (%) | 83 (100.0) | 403 (75.8) | <0.001 |
| HLBS total (29–145) | 104.1±6.5 | 116.8±14.7 | <0.001 |
| FSS mean (1–7) | 4.92±1.18 | 2.86±1.36 | <0.001 |
| Severe fatigue (FSS≥4), n (%) | 68 (81.9) | 111 (20.9) | <0.001 |
| IPAQ score | 22.6±14.8 | 28.4±16.2 | 0.003 |
| MSPSS (12–60) | 46.7±12.0 | 54.9±12.3 | <0.001 |
| Gestational age at birth (weeks) | 37.9±2.2 | 38.2±1.7 | 0.231 |
| Birth weight (g) | 3417±647 | 3288±484 | 0.081 |
| Apgar 1-min | 7.10±1.00 | 7.35±1.04 | 0.026 |
| Apgar 5-min | 8.64±0.60 | 8.70±0.62 | 0.276 |
| NICU admission, n (%) | 26 (31.3) | 60 (11.3) | <0.001 |
| Respiratory support, n (%) | 26 (31.3) | 60 (11.3) | <0.001 |

†Mann-Whitney U/$\chi^2$. Reported p-values are unadjusted; Benjamini-Hochberg FDR correction applied across the behavioral cluster (q<0.05 for all significant findings). BMI: body mass index; FSS: Fatigue Severity Scale; GDM: gestational diabetes mellitus; HLBS: Healthy Lifestyle Behaviors Scale; IPAQ: International Physical Activity Questionnaire; MSPSS: Multidimensional Scale of Perceived Social Support; NEQ: Night Eating Questionnaire; NES: night eating syndrome; NICU: neonatal intensive care unit; PSQI: Pittsburgh Sleep Quality Index.

vs. 2.86±1.36; d=1.55), with severe fatigue (FSS≥4) present in 81.9 vs. 20.9% (p<0.001). HLBS (d=0.97), physical activity (d=0.36), and MSPSS (d=0.67) scores were significantly lower in the NES-positive group.

Neonates of NES-positive mothers had lower 1-min Apgar scores (7.10±1.00 vs. 7.35±1.04; p=0.026), although gestational age at birth, birth weight, and 5-min Apgar scores did not differ. NICU admission was nearly threefold more common in the NES-positive group (31.3 vs. 11.3%; p<0.001; φ=0.21).

## Multivariable analyses

In Model A (Table 3), GDM was the strongest independent predictor of NES (adjusted odds ratio [AOR] 6.42; 95%CI 3.10–13.30; p<0.001). Fatigue (AOR 1.42), pre-pregnancy BMI (AOR 1.08), PSQI (AOR 1.07), and unplanned pregnancy (AOR 1.69) were also independently associated with NES, whereas higher HLBS (AOR 0.96) and physical activity (AOR 0.98) scores were protective. Model calibration (Hosmer-Lemeshow p=0.45) and discrimination (Nagelkerke $R^2$=0.31; classification accuracy 78.4%) were adequate.

In Model B, NES was independently associated with NICU admission (AOR 3.82; 95%CI 1.68–8.71; p=0.001) after adjustment for GDM status, BMI, sleep quality, fatigue, lifestyle behaviors, gestational age at birth, and maternal age. GDM (AOR 1.94; p=0.009) and gestational age at birth (AOR 0.71 per week; p<0.001) were also significant predictors. Model fit was acceptable (Nagelkerke $R^2$=0.22; Hosmer-Lemeshow p=0.38).

# DISCUSSION

This study demonstrates an elevated prevalence of NES among women with GDM (28.8%) compared with healthy pregnant controls (1.5%). NES clustered with poor sleep, severe fatigue, suboptimal lifestyle behaviors, lower physical activity, and reduced social support. To the best of our knowledge, this is among the first studies to concurrently characterize NES prevalence and its psychobehavioral correlates in a GDM cohort.

The 28.8% NES prevalence observed in our GDM cohort substantially exceeds the 2–6% commonly reported in general

**Table 3.** Multivariable logistic regression: independent predictors of (A) night eating syndrome and (B) neonatal intensive care unit admission.

| Predictor | AOR | 95%CI | p-value |
|---|---|---|---|
| **Model A—Outcome: NES (NEQ≥25)** | | | |
| GDM status (vs. control) | 6.42 | 3.10–13.30 | <0.001 |
| Pre-pregnancy BMI (per kg/m²) | 1.08 | 1.02–1.14 | 0.008 |
| PSQI score (per point) | 1.07 | 1.01–1.14 | 0.028 |
| FSS score (per point) | 1.42 | 1.18–1.71 | <0.001 |
| HLBS score (per point) | 0.96 | 0.93–0.99 | 0.009 |
| IPAQ score (per point) | 0.98 | 0.96–1.00 | 0.034 |
| Unplanned pregnancy (vs. planned) | 1.69 | 1.01–2.82 | 0.045 |
| Maternal age (per year) | 1.02 | 0.96–1.08 | 0.535 |
| Model fit: Nagelkerke $R^2$=0.31; Hosmer-Lemeshow p=0.45; accuracy=78.4% | | | |
| **Model B—Outcome: NICU admission** | | | |
| NES (NEQ≥25) | 3.82 | 1.68–8.71 | 0.001 |
| GDM status (vs. control) | 1.94 | 1.18–3.18 | 0.009 |
| Pre-pregnancy BMI (per kg/m²) | 1.05 | 1.00–1.10 | 0.062 |
| PSQI score (per point) | 1.03 | 0.99–1.07 | 0.140 |
| FSS score (per point) | 1.12 | 0.94–1.34 | 0.205 |
| HLBS score (per point) | 1.01 | 0.98–1.04 | 0.512 |
| Gestational age at birth (per week) | 0.71 | 0.60–0.84 | <0.001 |
| Maternal age (per year) | 0.99 | 0.94–1.04 | 0.621 |
| Model fit: events=86; Nagelkerke $R^2$=0.22; Hosmer-Lemeshow p=0.38; LR $\chi^2$=38.6, p<0.001 | | | |

AOR: adjusted odds ratio; BMI: body mass index; CI: confidence interval; FSS: Fatigue Severity Scale; GDM: gestational diabetes mellitus; HLBS: Healthy Lifestyle Behaviors Scale; IPAQ: International Physical Activity Questionnaire; LR: likelihood ratio; NEQ: Night Eating Questionnaire; NES: night eating syndrome; NICU: neonatal intensive care unit; PSQI: Pittsburgh Sleep Quality Index.

pregnant populations and approximates the high estimates documented among bariatric and severely obese cohorts[7]. Importantly, the increase in NES odds conferred by GDM persisted after adjustment for BMI, sleep, fatigue, and lifestyle factors, suggesting that GDM exerts an independent contribution beyond shared adiposity-related risk. This association may reflect circadian misalignment: altered glucose rhythms in GDM, combined with pregnancy-related shifts in melatonin secretion, cortisol patterning, and appetite-regulatory peptides, may amplify pre-existing vulnerabilities to disordered nocturnal eating[9,10]. The psychological burden of a GDM diagnosis—dietary restriction, glucose self-monitoring, and fetal-wellbeing anxiety—may further precipitate compensatory nocturnal intake[11].

The magnitude of psychobehavioral impairment in NES-positive women is noteworthy. Effect sizes for fatigue, sleep quality, and lifestyle behaviors were large by conventional criteria[21], indicating that NES rarely occurs in behavioral isolation. This pattern is consistent with the biobehavioral model proposed by Stunkard et al.[22], in which NES represents a circadian-linked disruption of sleep–wake and feeding–fasting cycles with downstream consequences for energy regulation, mood, and adaptive health behavior. The universally poor sleep quality and the higher rate of severe fatigue in NES-positive women plausibly compromise GDM self-management, including dietary adherence and glucose monitoring[4].

Higher HLBS and physical activity scores were independently protective against NES. These findings suggest that clinically meaningful behavioral improvements could translate into substantial reductions in NES risk. This observation is consistent with prior evidence that structured lifestyle interventions mitigate disordered eating and improve metabolic outcomes[23], supporting the integration of multidomain behavioral assessment into routine antenatal GDM care.

The independent association between NES and NICU admission represents one of the most clinically relevant findings. This relationship remained significant after adjustment for maternal covariates, suggesting that NES contributes to neonatal morbidity through pathways not fully explained by conventional metabolic risk factors. Possible mechanisms include impaired maternal glycemic control related to nocturnal caloric intake and circadian feeding misalignment[9], reduced treatment adherence potentially linked to fatigue and poor sleep[5], and altered placental nutrient transfer from caloric redistribution to nighttime hours—a possibility indirectly supported by the somewhat lower 1-min Apgar scores. Notably, gestational age at birth did not differ significantly between groups, suggesting that the higher rate of NICU admissions may reflect maternal behavioral and metabolic factors affecting neonatal transitional physiology rather than prematurity alone.

### Strengths and limitations

Strengths include a relatively large sample with a defined control group, and comprehensive validated psychometric assessment. To the best of our knowledge, this is among the first studies to concurrently characterize NES prevalence, its psychobehavioral correlates, and its independent association with neonatal outcomes in a well-defined GDM cohort. Several limitations warrant acknowledgment. First, the cross-sectional design precludes causal inference, and reverse causation cannot be excluded. Second, NES was ascertained by self-report rather than structured clinical interview, although the NEQ has well-established psychometric properties[15]. Third, the single-center setting may limit generalizability. Finally, data regarding HbA1c and continuous glucose monitoring parameters, detailed macronutrient intake, depressive symptomatology, and obstetric complications were unavailable, which may have influenced findings.

## CONCLUSION

NES affected nearly one in three women with GDM compared with only 1.5% of healthy pregnant controls, and GDM was independently associated with an increased likelihood of NES. NES clustered with universally poor sleep, severe fatigue, suboptimal lifestyle behaviors, reduced physical activity, and lower social support, and independently predicted an increased risk of neonatal morbidity. Healthy lifestyle behaviors and physical activity emerged as modifiable protective factors. These findings suggest that routine NEQ-based screening deserves further prospective evaluation to determine its utility in identifying NES as a candidate target for behavioral interventions aimed at improving maternal and neonatal outcomes.

## ETHICAL STATEMENT

Written informed consent was sought from each pregnant woman following an explanation of the purpose and design of the study. Pregnant women were informed that their involvement in the study was voluntary and that they could withdraw at any time. In addition, they were given the assurance that their data would be kept confidential. This study followed the ethical principles outlined in the Declaration of Helsinki and was approved by the Research Ethics Committee (Approval No. 73; dated May 10, 2023).

## AUTHORS' CONTRIBUTIONS

**EK:** Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Supervision, Writing – original draft, Writing – review and editing. **LK:** Data curation, Investigation, Resources, Writing – original draft, Writing – review and editing. **CEÖ:** Formal Analysis, Investigation, Writing – original draft, Writing – review and editing. **ZK:** Data curation, Investigation, Resources, Writing – original draft, Writing – review and editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. American Diabetes Association Professional Practice Committee. Classification and diagnosis of diabetes: standards of care in diabetes—2024. Diabetes Care. 2024;47(Suppl 1):S2042. https://doi.org/10.2337/dc24-S002
2. Ejaz Z, Azhar Khan A, Sebghat Ullah S, Aamir Hayat M, Maqbool MA, Amin Baig A. The effects of gestational diabetes on fetus: a surveillance study. Cureus. 2023;15(2):e35103. https://doi.org/10.7759/cureus.35103
3. Umer A, Kanwal S, Ahmad A, Fatima A, Fatima A, Farooq F, et al. Maternal and neonatal outcomes associated with preeclampsia and gestational diabetes mellitus: a retrospective cohort study. Cureus. 2025;17(6):e86120. https://doi.org/10.7759/cureus.86120
4. Faal Siahkal S, Javadifar N, Najafian M, Iravani M, Zakerkish M, Heshmati R. The psychosocial challenges associated with gestational diabetes mellitus: a systematic review of qualitative studies. Prim Care Diabetes. 2022;16(1):11-26. https://doi.org/10.1016/j.pcd.2021.09.003
5. Lavery ME, Frum-Vassallo D. An updated review of night eating syndrome: an under-represented eating disorder. Curr Obes Rep. 2022;11(4):395-404. https://doi.org/10.1007/s13679-022-00487-9
6. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Arlington, VA: American Psychiatric Publishing; 2013.
7. Allison KC, Crow SJ, Reeves RR, West DS, Foreyt JP, Dilillo VG, et al. Binge eating disorder and night eating syndrome in adults with type 2 diabetes. Obesity (Silver Spring). 2007;15(5):1287-93. https://doi.org/10.1038/oby.2007.150
8. Tzischinsky O, Latzer Y. Night eating syndrome prevalence and its association with sleep quality, eating patterns, and psychopathology in an Israeli community sample. Soc Psychiatry Psychiatr Epidemiol. 2025;60(12):2859-69. https://doi.org/10.1007/s00127-025-02880-w
9. Bargagna M, Casu M. Night eating syndrome: a review of etiology, assessment, and suggestions for clinical treatment. Psychiatry Int. 2024;5(2):289-304. https://doi.org/10.3390/psychiatryint5020020
10. Brown RB. Late-night feeding, sleep disturbance, and nocturnal congestion mediated by hyperglycemia, renal sodium retention, and cortisol: a narrative review. Clocks Sleep. 2025;8(1):1. https://doi.org/10.3390/clockssleep8010001
11. Gibė E, Perminas A. Psychological, social, and sociocultural factors related to disordered eating behaviour during pregnancy: a systematic review. J Health Psychol. 2026;13591053261442874. https://doi.org/10.1177/13591053261442874
12. Bannatyne AJ, Hughes R, Stapleton P, Watt B, MacKenzie-Shalders K. Signs and symptoms of disordered eating in pregnancy: a Delphi consensus study. BMC Pregnancy Childbirth. 2018;18(1):262. https://doi.org/10.1186/s12884-018-1849-3
13. Zhu S, Surampudi P, Field NT, Chondronikola M. Meal timing and glycemic control during pregnancy—is there a link? Nutrients. 2021;13(10):3379. https://doi.org/10.3390/nu13103379
14. Loy SL, Ku CW, Zheng RT, Lim CHF, Chang TY, Chen LW, et al. Associations of predominant night-eating with plasma glycemic status and continuous glucose monitoring measures among pregnant women. Clin Nutr. 2023;42(12):2320-7. https://doi.org/10.1016/j.clnu.2023.10.009
15. Allison KC, Lundgren JD, O'Reardon JP, Martino NS, Sarwer DB, Wadden TA, et al. The Night Eating Questionnaire (NEQ): psychometric properties of a measure of severity of the Night Eating Syndrome. Eat Behav. 2008;9(1):62-72. https://doi.org/10.1016/j.eatbeh.2007.03.007
16. Buysse DJ, Reynolds CF 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res. 1989;28(2):193-213. https://doi.org/10.1016/0165-1781(89)90047-4
17. Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. Arch Neurol. 1989;46(10):1121-3. https://doi.org/10.1001/archneur.1989.00520460115022
18. Walker SN, Sechrist KR, Pender NJ. The health-promoting lifestyle profile: development and psychometric characteristics. Nurs Res. 1987;36(2):76-81.
19. Zimet GD, Dahlem NW, Zimet SG, Farley, GK. The multidimensional scale of perceived social support. J Pers Assess. 1988;52(1):30-41.
20. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc. 2003;35(8):1381-95. https://doi.org/10.1249/01.MSS.0000078924.61453.FB
21. Elsahoryi NA, Ibrahim MO, Hammouh F, Alhaj OA, Al-Basha S. Unveiling night eating syndrome: how it connects to mental health, insomnia, and quality of life in university students-a cross-sectional study. J Eat Disord. 2025;13(1):225. https://doi.org/10.1186/s40337-025-01406-8
22. Stunkard AJ, Allison KC, Lundgren JD, O'Reardon JP. A biobehavioural model of the night eating syndrome. Obes Rev. 2009;10(Suppl 2):69-77. https://doi.org/10.1111/j.1467-789X.2009.00668.x
23. Sparks JR, Ghildayal N, Hivert MF, Redman LM. Lifestyle interventions in pregnancy targeting GDM prevention: looking ahead to precision medicine. Diabetologia. 2022;65(11):1814-24. https://doi.org/10.1007/s00125-022-05658-w

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260803

# Rational Drug Use Support System: design, artificial intelligence integration, and pilot feasibility trial

Mustafa Kahraman[1], Meris Esra Bozkurt[2]*, Soner Kotan[3], Buse Nur Ozturk[4], Rana Aydın[4], Saim Hatipoğlu[5], Hasan Ali Eris[6], Mehmet Ali Erturk[7], Muhammed Ali Aydin[6], Mehmet Akif Karan[8]

## SUMMARY

**OBJECTIVE:** The Rational Drug Use Support System was developed to assist patients in using the right drug, at the right time and at the right dose according to their treatment protocols. The aim of this study was to evaluate the performance of an artificial intelligence-enhanced mobile application designed to support rational drug use and to assess its feasibility and usability through a pilot trial.

**METHODS:** The system includes an artificial intelligence -powered mobile application and a web-based interface integrating convolutional neural networks and object detection algorithms for human activity recognition. To build an image library, real patient videos were recorded in a hospital setting under physician supervision. A total of 357 original medication-intake videos were collected from 31 patients, which were then segmented into 1,626 short clips to train the artificial intelligence model. The system was first tested under controlled laboratory conditions and then evaluated for feasibility in a hospital pilot trial involving eight participants over 15 days.

**RESULTS:** The proposed artificial intelligence model achieved 94% accuracy in recognizing medication intake activities in the test set, suggesting promising performance that warrants further evaluation. Exploratory analyses of the 36-Item Short Form Health Survey Questionnaire administered to patients who utilized the mobile application for 15 days revealed improvements in the domains of social functioning, vitality, and mental health.

**CONCLUSION:** The findings suggest that the "Rational Drug Use Support System" appears feasible for implementation in telemedicine and clinical settings to support rational drug use and patient engagement in chronic disease management. However, larger controlled studies are warranted to confirm these preliminary findings.

**KEYWORDS:** Digital health. Drug adherence. Telemedicine. Artificial intelligence. Mobile health.

## INTRODUCTION

Chronic diseases remain a major global health challenge and are responsible for a substantial share of mortality, years of life lost, and long-term morbidity worldwide[1,2]. A major barrier to effective chronic disease management is rational drug use and deviation from prescribed treatment protocols. The World Health Organization has long emphasized that medication-related problems are common and that many patients do not take medicines as prescribed[3]. Poor adherence can reduce treatment effectiveness, increase preventable complications and hospitalizations, and, particularly in antimicrobial use, contribute to drug resistance. In practice, inappropriate drug use may include missed doses, delays, or self-directed changes in dose or treatment duration, all of which can lead to substantial clinical and economic consequences.

Digital health interventions, particularly mobile health (mHealth) applications, have therefore gained attention as

[1]Türkiye Public Health and Chronic Diseases Institute – Istanbul, Türkiye.
[2]Mersin University, Mersin Medical School, Department of Internal Medicine, Division of Geriatrics – Mersin, Türkiye.
[3]Türkiye Health Data Research and Artificial Intelligence Applications Institute – Istanbul, Türkiye.
[4]Istanbul University, Istanbul Faculty of Medicine – Istanbul, Türkiye.
[5]Istanbul Gelisim University, Faculty of Engineering and Architecture, Department of Software Engineering – Istanbul, Türkiye.
[6]Istanbul University-Cerrahpasa, Department of Computer Engineering – Istanbul, Türkiye.
[7]Istanbul University, Department of Computer Engineering – Istanbul, Türkiye.
[8]Istanbul University, Istanbul Faculty of Medicine, Department of Internal Medicine, Division of Geriatrics – Istanbul, Türkiye.
*Corresponding author: mesragunduz@gmail.com

Conflicts of interest: the authors declare there is no conflicts of interest. Funding: this study was funded by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) Project no. 7218037. Except for this, this research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
Received on April 13, 2026. Accepted on July 19, 2026.
Scientific Editor: José Maria Soares Júnior

scalable tools to support self-management, especially in settings where smartphone use is widespread[4]. In this context, digitally supported care models such as telehealth and telemedicine have become increasingly relevant for maintaining continuity of care, supporting rational drug use, and reducing avoidable healthcare use[5,6].

In this study, we developed an artificialintelligence (AI)-driven Rational Drug Use Support System to support patients in taking the right medication at the right time and dose. The system was developed and evaluated in a real hospital setting under clinical supervision. The study aimed to assess the performance of the AI-enabled system in objectively verifying medication intake activities and to evaluate its feasibility and usability in routine clinical practice through a pilot trial.

## METHODS

### Study design and participants

This study used a multi-stage design to develop and clinically evaluate a Rational Drug Use Support System. The clinical component was designed as a feasibility pilot study rather than an efficacy trial. A preliminary survey of 300 patients was conducted to examine factors related to rational drug use and technology use among individuals with chronic diseases[7]. Of these, 31 volunteered to contribute to image library development, and eight were subsequently selected by computer-generated randomization for inclusion in the hospital-based pilot trial. Eligible participants were adults aged 18 years or older who were receiving continuous medication, able to swallow and self-administer medication independently, capable of providing written informed consent, and had basic technological literacy. Individuals with cognitive impairment, severe sensory or functional limitations affecting independent drug use, severe visual impairment, life expectancy under two years, pregnancy or breastfeeding, or inability to self-administer medication were excluded.

### Image library development and artificial intelligence model training

Video data were collected in the inpatient unit of a university internal medicine department, under physician supervision and in accordance with privacy safeguards. Before recording, demographic characteristics, prescribed medication count, and general health status were documented, and body mass index (BMI) was calculated as weight in kilograms divided by height in meters squared[8]. To increase variability in the training data, participants were recorded while sitting, standing, and lying in bed. Videos were captured using three smartphones mounted on tripods and positioned to provide a clear view of the face and upper body. Each recording followed a two-step procedure in which the participant first showed the medication package or blister pack to the camera and then completed the medication-taking action. For each participant, 10–14 recordings were obtained under varying hand-use, lighting, and background conditions. Individual recordings lasted 2–6 s, yielding 357 original videos from 31 participants. All videos were anonymized and securely transferred for AI model development.

For model development, the 357 original recordings were randomly partitioned into training, validation, and test sets at a ratio of 60:20:20. The training set was used for model parameter optimization, the validation set was used for hyperparameter tuning and early stopping to prevent overfitting, and the test set was reserved exclusively for final performance evaluation.

From these 357 recordings, 1,626 short clips of 2–6 s were generated to capture different stages of medication intake and increase within-class variability. To improve generalizability and reduce confusion with visually similar actions, the dataset was expanded with additional eating clips from publicly available sources and the UCF101 (A Dataset of 101 Human Action Classes From Videos in the Wild) Human Activity Recognition dataset, together with clips from 11 unrelated activities used as negative examples.

Medication-intake actions were detected using the Personalized Human Activity Recognition (PHAR) framework[9], which combines object detection with human activity recognition (HAR) to focus on regions most relevant to the target behavior. Two PHAR-based architectures were implemented in line with previous HAR studies: a Convolutional Neural Network Long Short-Term Memory (CNN–LSTM) model and a 3D convolutional neural network (3D-CNN)[7,,9,10,11]. These were developed as two model variants, Two-Dimensional Motion Feature Maps (MFM2D) and Three-Dimensional Motion Feature Maps (MFM3D), which shared the same input structure and preprocessing pipeline but differed in temporal modelling.

The PHAR input consisted of 16 temporally ordered frames, referred to as the Main Frame sequence (MFRM). Because medication-taking and drinking actions mainly involve coordinated movements of the face, hands, and upper body, object detection was used to generate three additional focus views for each frame: Face Frame (FFRM), Hands Frame (HFRM), and Person Frame (PFRM).

The validation set was used to monitor training and limit overfitting, whereas the test set was reserved for final performance evaluation. Because the dataset contained a substantial number of negative examples, model performance was assessed

not only by accuracy but also by precision, recall, and F1-score for the three target classes: Taking Medicine, Drinking Water, and Eating. The F1-score was considered the primary balanced performance metric.

### System architecture and functionality

The Rational Drug Use Support System was implemented as a client–server architecture comprising a physician-facing web platform, a patient-facing mobile application, and a central server hosting the trained PHAR model. After authentication, it synchronized the patient's medication schedule, delivered reminders at predefined times, and prompted the user to record a short video of the medication intake process. These data were securely transmitted to the server, where the PHAR model classified the action and generated a medication intake outcome.

### Pilot feasibility study

Feasibility evaluation was conducted through a prospective observational pilot study involving eight participants selected from the video dataset group. The study period was 15 days. The primary focus of this pilot was to assess feasibility, usability, and user acceptance rather than clinical efficacy. Key feasibility outcomes included successful video uploads, system usability, application adherence, and medication verification success rates. At baseline, all participants received standardized training on application use and alert messages. Physicians created user accounts through the web platform and entered the prescribed treatment schedules, which were then synchronized to participants' smartphones. Secondary exploratory outcomes were assessed using the 36-Item Short Form Health Survey (SF-36) Short Form Health Survey at baseline (day 0) and post-intervention (day 15), together with a study-specific patient satisfaction questionnaire[12-14].

### Statistical analysis

All data were recorded and analyzed using Statistical Package for the Social Sciences (SPSS) Statistics (version 25.0; IBM/SPSS). Analyses were performed separately for the video dataset group and the clinical pilot group. Normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Normally distributed variables are presented as mean±standard deviation, and non-normally distributed variables as median (min–max). Associations between categorical variables were examined using the chi-square test. For the pilot study, SF-36 scores at baseline and day 15 were compared using paired analyses, with the paired t-test applied for normally distributed differences and the Wilcoxon signed-rank test used otherwise. Statistical significance was set at a two-sided $p<0.05$.

## RESULTS

### Participant characteristics

The image library cohort comprised 31 participants with a mean age of 44±18 years, of whom 67.7% were female. Polypharmacy, defined as the use of four or more medications, was observed in 6.5% of participants. Demographic and clinical characteristics of this cohort are presented in Table 1.

The pilot study included eight participants with a mean age of 38 years (range: 30–53) and a mean BMI of 25.3±3.3 kg/m². Several indicators pointed to a need for medication intake support: 50% reported frequently forgetting medication intake, 25% reported receiving assistance for drug use, and 87.5% reported both discarding 1–3 boxes of medication annually due to expiry and keeping unfinished medication boxes at home. Before using the system, 75% indicated that they expected the application to be beneficial. Participant characteristics for the pilot group are summarized in Table 2. Within this small sample, gender differences were observed for BMI ($p=0.002$) and living situation ($p=0.028$).

### Artificial intelligence model performance

The proposed AI model achieved 94% accuracy on the test set in recognizing medication intake activities. The MFM3D variant demonstrated superior performance compared to the MFM2D variant, particularly for the "Taking Medicine" activity class. Precision, recall, and F1-scores for the target classes were consistently high (>0.90), indicating robust classification performance.

### Pilot feasibility outcomes

With regard to feasibility outcomes, all eight participants successfully completed the 15-day study period. The application adherence rate was high, with participants recording videos for the majority of their scheduled medication intake times. Successful video upload rates exceeded 90%, and the system demonstrated reliable medication verification. System usability was rated favorably by participants.

### Exploratory health-related quality of life outcomes

Health-related quality of life was assessed using the SF-36 at baseline and after 15 days of application use. General health scores did not change significantly (70.6±17.4 vs. 71.8±12.5; $p=0.104$). Exploratory analyses revealed statistically significant improvements in vitality/energy ($p=0.001$), mental health/emotional well-being ($p=0.031$), and social functioning ($p=0.019$), whereas no significant differences were found in the remaining domains (Table 2).

**Table 1.** Basic data of participants who used to mobile application by gender.

| | Male (n=3) 37.5% | Female (n=5) 62.5% | Total (n=8) 100% | p-value |
|---|---|---|---|---|
| Age (year)* | 23 (20–53) | 20 (20–49) | 28 (20–53) | 0.341 |
| BMI (kg/m²) | 28.8+0.9 | 22.9+1.7 | 25.1+3.3 | 0.002[Ø] |
| Marital status (n, %) | | | | |
| Married | 1 (33.3%) | 1 (20.0%) | 2 (25.0%) | 0.673 |
| Single | 2 (66.7%) | 5 (80.0%) | 6 (75.0%) | |
| Living situation (n, %) | | | | |
| Family | 0 (0.0%) | 2 (40.0%) | 5 (62.5%) | 0.028[Ø] |
| Alone | 3 (100%) | 3 (60.0%) | 3 (37.5%) | |
| Education (n, %) | | | | |
| High school | 0 (0.0%) | 1 (20.0%) | 1 (12.5%) | 0.408 |
| College-faculty | 3 (100%) | 4 (80.0%) | 6 (87.5%) | |
| Profession (n, %) | | | | |
| Housewife | 0 (0.0%) | 1 (20.0%) | 6 (75.0%) | 0.315 |
| Studen t | 2 (66.7%) | 4 (80.0%) | 1 (12.5%) | |
| Teacher | 1 (33.3%) | 0 (0.0%) | 1 (12.5%) | |
| Income level (n, %) | | | | |
| Minimum wage | 0 (0%) | 0 (0%) | 1 (12.5%) | 0.315 |
| Between minimum wage and twice | 2 (66.7%) | 1 (20%) | 6 (75.0%) | |
| Minimum wage is between 2 and 4 times | 1 (33.3%) | 4 (80%) | 1 (12.5%) | |
| Chronic medication using status (n, %) | 2 (66.7%) | 1 (20%) | 3 (37.5%) | 0.187 |
| Time to take medication | | | | |
| Morning | 1 (33.3%) | 4 (80.0%) | 5 (62.5%) | 0.293 |
| Evening | 1 (33.3%) | 0 (0.0%) | 1 (12.5%) | |
| Noon | 1 (33.3%) | 1 (20.0%) | 2 (25.0%) | |
| The state of forgetting the medication time (n, %) | 2 (66.7%) | 3 (60.0%) | 5 (62.5%) | 0.850 |
| Frequency of forgetting the medication time (n, %) | | | | |
| Often | 2 (66.7%) | 2 (40.0%) | 4 (50.0%) | 0.641 |
| Sometimes | 0 (0.0%) | 1 (20.0%) | 1 (12.5%) | |
| Never | 1 (33. 3%) | 2 (40.0%) | 3 (37.5%) | |
| The state of forgetting the medication according to the hungry or full state (n, %) | | | | |
| Often | 0 (0.0%) | 1 (20.0%) | 1 (12.5%) | 0.408 |
| Never | 3 (100%) | 4 (80.0%) | 0 (0.0%) | |
| Presence of situation of discarding medication because it has expired (n, %) | 3 (100%) | 4 (80.0%) | 7 (87.5%) | 0.408 |
| Number of boxes of drugs discarded because the expiration date has passed (n, %) | | | | |
| 0 box | 0 (0.0%) | 1 (20.0%) | 0 (0.0%) | 0.408 |
| 1–3 boxes | 3 (100%) | 4 (80.0%) | 7 (87.5%) | |
| 4–7 boxes | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | |
| Presence of unfinished medication in home (n, %) | 3 (100%) | 4 (80%) | 7 (87.5%) | 0.408 |
| Number of unfinished medicine boxes at home (n, %) | | | | |
| 0 box | 0 (0.0%) | 1 (20.0%) | 0 (0.0%) | 0.376 |
| 1–3 boxes | 1 (33.3%) | 3 (60.0%) | 4 (50.0%) | |
| 4–7 boxes | 2 (66.7%) | 1 (20.0%) | 3 (37.5%) | |
| Situation of getting help with medication using (n, %) | 1 (33.3%) | 1 (20.0%) | 2 (25.0%) | 0.673 |
| Knowing the “e- nabiz” application (n, %) | 2 (66.7%) | 5 (100%) | 7 (87.5%) | 0.168 |
| Opinion about the “e-nabiz” application (n, %) | | | | |
| Usefu l | 2 (66.7%) | 4 (80.0%) | 6 (75.0%) | 0.673 |
| Undecided | 1 (33.3%) | 1 (20.0%) | 2 (25.0%) | |
| Opinion about the rational drug use support system mobile application (n, %) | | | | |
| Useful | 1 (33.3%) | 5 (100%) | 6 (75.0%) | 0.035[Ø] |
| Undecided | 2 (66.7%) | 0 (0.0%) | 2 (25.0%) | |

m: meter; kg: kilogram; BMI: body mass index. Data are given as mean±standard deviation, median (interquartile range) or number (percentage) as applicable. *Given data as median (interquartile range). [Ø]Significant p-values.

### User satisfaction

At the end of the pilot, participants reported high satisfaction with the system. Overall, 75% stated that the application supported rational drug use of their medication regimen, while 87.5% reported that reminders were effective and that the system helped them manage drug use independently. Interoperability was also identified as a key priority, with 87.5% indicating that integration with public and private hospitals, pharmacies, and the national e-Nabiz ® personal health system would be useful (Table 3).

## DISCUSSION

This study describes the development and preliminary feasibility evaluation of an AI-enabled Rational Drug Use Support System that integrates a PHAR pipeline for the objective verification of medication intake. The primary aim was to assess the performance of the AI system and evaluate its feasibility and usability in a clinical setting, rather than to establish clinical effectiveness. Two main findings emerged. First, the proposed PHAR variants (MFM2D and MFM3D) outperformed the baseline architectures used in this study, particularly for the "Taking Medicine" activity. Second, in the short hospital-based pilot, participants demonstrated high engagement with the system and reported favorable usability, together with exploratory improvements in selected SF-36 domains.

Previous studies on medication intake monitoring have reported a wide range of classification performance across image- and sensor-based systems. Scoping evidence suggests that accuracies often remain below the mid-90% range[15], although some multimodal approaches have achieved higher values under specific conditions[16]. The performance observed in the present study is broadly consistent with this literature and supports the feasibility of vision-based approaches for objective intake verification in controlled settings. Importantly, reducing false positives is likely to be critical for sustained use, as frequent incorrect alerts or non-adherence flags may increase user burden and contribute to disengagement over time[17]. From this perspective, combining personalized attention to relevant body regions with temporal modelling may improve the reliability

**Table 2.** Comparison of the general quality of life of the participants using the mobile application, evaluated by the 36-item Short Form Health Survey scale before and after the application using.

| | SF-36 scale before using application | SF-36 scale after using application | p-value |
|---|---|---|---|
| 1. Physical functioning (%) | 93.7±11.5 | 93.7±9.1 | 1 |
| 2. Role limitations due to physical health (%) | 84.3±29.6 | 90.6±26.5 | 0.32 |
| 3. Role limitations due to emotional problems (%) | 29.1±33.3 | 54.1±50.2 | 0.08 |
| 4. Energy/fatigue (%) | 44.3±24.9 | 51.2±22.3 | 0.001* |
| 5. Emotional well-being (%) | 61.0±18.9 | 67.0±16.2 | 0.03* |
| 6. Social functioning (%) | 78.1±24.7 | 81.2±25.0 | 0.02* |
| 7. Pain (%) | 83.7±15.6 | 80.9±16.3 | 0.46 |
| 8. General health (%) | 70.6±17.4 | 71.8±12.5 | 0.10 |

SF-36: The 36-item Short Form Health Survey. Data are given as mean±standard deviation, median (interquartile range) or number (percentage) as applicable. The '*' symbol represents significant p-values.

**Table 3.** Satisfaction survey of users of the Rational Drug Use Support System mobile application regarding the pilot trial.

| | Opinion useful group | Opinion non useful group |
|---|---|---|
| Did the mobile application help you comply by the medication treatment protocol? | 6 (75.0%) | 2 (25.0%) |
| Did the mobile application successfully perform your medication time reminders? | 7 (87.5%) | 1 (12.5%) |
| Did the mobile application successfully perform fasting/satiety reminders for medication use? | 6 (75.0%) | 2 (25.0%) |
| Were you able to use your medicines without any help from anyone with the support of the mobile application? | 7 (87.5%) | 1 (12.5%) |
| Did your medication using regulated in your daily life thanks to the use of the mobile application? | 6 (75.0%) | 2 (25.0%) |
| Are you satisfied with using the mobile app? | 6 (75.0%) | 2 (25.0%) |
| Would you like the mobile application to work integrated with all public/private hospitals? | 7 (87.5%) | 1 (12.5%) |

of AI-driven monitoring, particularly when activities are visually confusable.

Although the pilot was not designed to establish clinical efficacy, participant-reported outcomes showed exploratory improvements over the 15-day period in vitality, mental health, and social functioning. The direction of change observed in this pilot is also consistent with broader evidence suggesting that mHealth interventions can support self-management and may improve patient-reported outcomes in chronic illness[18]. However, given the exploratory design of the study, the absence of a control group, and the small sample size, these findings should be interpreted as preliminary and hypothesis-generating rather than as evidence of clinical effectiveness. It is possible that the observed improvements reflect regression to the mean, placebo effects, or increased attention from study participation rather than the intervention itself. These findings require confirmation in larger and more rigorous evaluations with control groups and longer follow-up periods.

User feedback also supported the feasibility of the system. Most participants reported that reminders were helpful and that the system enabled them to manage drug use more independently, in line with earlier findings on AI-assisted medication management tools[19]. Participants also emphasized the importance of interoperability, particularly integration with hospitals, pharmacies, and the national personal health record platform, e-Nabiz ®[20]. Such integration could simplify medication scheduling, reduce manual data entry, and support continuity across care settings, although it would also require careful attention to governance, authentication, and data protection.

A major strength of this study is the end-to-end development and evaluation of an integrated medication intake support system using real patient videos collected in a clinical environment. This setting enhances the ecological validity of both the model-development pipeline and the usability assessment. In addition, combining patient-reported Health-Related Quality of Life (HRQoL) (SF-36) with satisfaction feedback provides a broader perspective on both experiential and outcome-related dimensions.

Several limitations should be acknowledged. The pilot included a small sample (n=8) and a short follow-up period of 15 days, which limits generalizability and precludes firm conclusions about clinical effectiveness. The absence of a control group further limits the interpretation of the observed SF-36 changes. In addition, the study population was relatively young, and the proportion of participants with polypharmacy was low, which may limit the applicability of the findings to older adults and patients with cognitive impairment who are at highest risk for medication-related problems. The HRQoL changes observed should therefore be regarded as exploratory and hypothesis-generating.

Future research should include larger, controlled trials with longer follow-up periods to evaluate the clinical effectiveness of the system. Studies should also specifically target older adult populations and patients with polypharmacy, who stand to benefit most from such interventions. Additionally, objective measures of medication adherence should be incorporated alongside patient-reported outcomes to provide a more comprehensive assessment of the system's impact.

## CONCLUSION

This study introduced and preliminarily evaluated an AI-enabled Rational Drug Use Support System designed to support rational drug use via automated activity recognition within a mobile, patient-facing workflow. The findings suggest that the system appears feasible for implementation in clinical settings, with high AI classification accuracy and favorable user satisfaction reported in a small pilot sample. However, the observed improvements in HRQoL domains should be interpreted as exploratory, and larger controlled studies are needed to confirm these preliminary findings and establish clinical effectiveness. Overall, the findings support the feasibility of combining mobile health delivery with AI-based verification as a complementary approach to conventional chronic disease management.

## ETHICAL APPROVAL

Ethical approval for the study was obtained from the Istanbul University Faculty of Medicine Clinical Research Ethics Committee (Approval No: 10.12.2021/1895). Participants in this clinical pilot trial were provided with both oral and written information about the study and gave their written informed consent for participation and publication.

## AUTHORS' CONTRIBUTIONS

MK: Conceptualization, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MEB: Data curation, Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SK: Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. BNO: Data curation, Formal

Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. RA: Data curation, Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. SH: Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. HAE: Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MAE: Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MAA: Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. MAK: Formal Analysis, Investigation, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. GBD 2016 Causes of Death Collaborators. Global, regional, and national age-sex specific mortality for 264 causes of death, 1980–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet. 2017;390(10100):1151-210. https://doi.org/10.1016/S0140-6736(17)32152-9
2. GBD 2017 Causes of Death Collaborators. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2018;392(10159):1736-88. https://doi.org/10.1016/S0140-6736(18)32203-7
3. World Health Organization. Adherence to long-term therapies: evidence for action. World Health Organization; 2003. [cited on 2025 Nov 10]. Available from: https://iris.who.int/handle/10665/42682
4. Kahraman, M, Çakmak R, Satman I, Karşıdağ K, Öztürk Ş, Çelik MÖ, et al. The effect of smartphone apps and technology compatibility on diabetes control in diabetic patients using insulin. J Istanbul Fac Med. 2021;84(4):543-53. https://doi.org/10.26650/IUITFD.2021.869948
5. Chudasama YV, Gillies CL, Zaccardi F, Coles B, Davies MJ, Seidu S, et al. Impact of COVID-19 on routine care for chronic diseases: a global survey of views from healthcare professionals. Diabetes Metab Syndr. 2020;14(5):965-7. https://doi.org/10.1016/j.dsx.2020.06.042
6. Mashoudy KD, Perez SM, Nouri K. From diagnosis to intervention: a review of telemedicine's role in skin cancer care. Arch Dermatol Res. 2024;2;316(5):139. https://doi.org/10.1007/s00403-024-02884-7
7. Girshick R, Donahue J, Darrell T, Malik J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In proceeding of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, 2014, p. 580-7. https://doi.org/10.1109/CVPR.2014.81
8. World Health Organization. Obesity and overweight. 2025. [cited on 2024 Sep 4]. Available from: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight
9. Eriş HA, Ertürk MA, Aydın MA. A framework for personalized human activity recognition. Int J Pattern Recognit Artif Intell. 2023;37(10):2356016. https://doi.org/10.1142/s0218001423560165
10. Ferner RE, Aronson JK. Communicating information about drug safety. BMJ 2006;333:143-5.
11. Ullah A, Ahmad J, Muhammad K, Sajjad M, Baik SW. Action recognition in video sequences using deep bi-directional LSTM with CNN features. IEEE Access; 2018;6:1155-66. https://doi.org/10.1109/ACCESS.2017.2778011
12. Mithe R, Supriya I, Nilam D. Optical character recognition. Int J Recent Technol Eng (IJRTE). 2013;2(1):72-5.
13. Redmon J, Divvala S, Girshick R, Farhad A. You only look once: unified, real-time object detection. IEEE. 2016;779-88. https://doi.org/10.1109/CVPR.2016.91
14. Kocyigit H. Kısa form-36 (kf-36)'nın Türkçe versiyonunun güvenilirliği ve geçerliliği. İlaç ve Tedavi Dergisi. 1999;12(1):102-6.
15. Bohlmann A, Mostafa J, Kumar M. Machine learning and medication adherence: scoping review. JMIRx Med. 2021;2(4): e26993.
16. Uppuluri V. Real-time monitoring of patient adherence using AI. Int J Health Sci. 2025;8(3):52-68.
17. Pulice E, Coustasse A. AI-driven solutions promote medication adherence. Pharmacy Times; 2024. p. 42-4.
18. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? NPJ Digit Med. 2020;3(1):4. https://doi.org/10.1038/s41746-019-0206-x
19. Republic of Türkiye Ministry of Health. E-Nabiz personal health record system. 2024. [cited on 2025 Nov 19]. Available from: https://enabiz.gov.tr
20. Moreno-Ligero M, Moral-Munoz JA, Salazar A, Failde I. mHealth intervention for improving pain, quality of life, and functional disability in patients with chronic pain: systematic review. JMIR mHealth uHealth. 2023;11(1):e40844. https://doi.org/10.2196/40844

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260842

# Association of Vascular Endothelial Growth Factor A-2549 I/D polymorphism with polycystic ovary syndrome susceptibility in Turkish women

Recai Aci[1]*, Mehmet Can Nacar[2], Serbulent Yigit[3], Husniye Rustemoglu[4], Eda Nacar[5]

## SUMMARY

**OBJECTIVE:** The aim of this case-control study was to evaluate whether the Vascular Endothelial Growth Factor A promoter rs35569394 (-2549 insertion/deletion) polymorphism is associated with polycystic ovary syndrome risk and related clinical-metabolic features.

**METHODS:** The study included 104 women with polycystic ovary syndrome diagnosed according to Rotterdam criteria and 128 controls. Clinical, hormonal, and metabolic variables were recorded, homeostatic model assessment of insulin resistance was calculated, and Vascular Endothelial Growth Factor A-2549 insertion/deletion genotypes were determined by direct polymerase chain reaction. Hardy-Weinberg equilibrium was assessed in controls. Benjamini-Hochberg false-discovery-rate-adjusted p-values were calculated for the 11 clinical comparisons, while prespecified Bonferroni-adjusted p-values were retained for genotype and family-history contrasts.

**RESULTS:** Body mass index, prolactin, homeostatic model assessment of insulin resistance, insulin, follicle stimulating hormone, luteinizing harmone, and total testosterone were higher in the polycystic ovary syndrome group. Overall genotype distribution differed between groups (Cramer's V=0.20; chi-square=8.87, p=0.012). The detection/deletion genotype was more frequent in polycystic ovary syndrome patients than controls (30.77% vs. 15.63%; Cohen's h=0.36; raw p=0.007; Bonferroni-adjusted p=0.028; OR 2.40, 95%CI 1.28–4.51). The family-history detection/deletion comparison did not remain significant after adjustment (raw p=0.018; adjusted p=0.072). Hardy-Weinberg equilibrium analysis in controls yielded an exact p-value of 0.050 and a Pearson p-value of 0.037.

**CONCLUSION:** In this Turkish case-control sample, the Vascular Endothelial Growth Factor A-2549 insertion/deletion detection/deletion genotype was more frequent among women with polycystic ovary syndrome than among controls. This observed association does not establish causality, a functional effect, phenotype-specific relevance, or clinical utility. Deviation from Hardy-Weinberg equilibrium in controls, residual confounding from incomplete matching, limited power, and lack of functional validation make the finding preliminary. Independent validation in larger, well-matched cohorts with prespecified genotyping quality control, phenotype stratification, and functional studies is required.

**KEYWORDS:** Polycystic ovary syndrome. Vascular Endothelial Growth Factor A. Polymorphism, genetic. Insulin resistance.Genotype.

## INTRODUCTION

Polycystic ovary syndrome (PCOS) is a common endocrine-metabolic disorder characterized by ovulatory dysfunction, hyperandrogenism and/or polycystic ovarian morphology[1,2]. Its heterogeneous expression is influenced by adiposity, insulin resistance, gonadotropin imbalance and genetic background. Prevalence varies by diagnostic criteria and ethnicity, supporting the study of population-specific genetic markers[3-5]. The contribution of angiogenesis-related variants remains incompletely defined.

The promoter variant rs35569394 (-2549 insertion/deletion (I/D)) is an 18-bp insertion/deletion located in the VEGFA promoter and has been hypothesized to influence transcriptional regulation; however, its functional direction and magnitude in ovarian tissue remain uncertain.

Associations between VEGFA polymorphisms and PCOS vary across populations, possibly because of ethnic, linkage-disequilibrium and environmental differences[6-13]. Evidence regarding rs35569394 in PCOS remains limited and population dependent, and the variant has not been adequately characterized in Turkish women with PCOS. Familial aggregation makes family history a clinically relevant, although imprecise, marker of inherited susceptibility. We therefore hypothesized

[1]Aydin Adnan Menderes University, Soke Vocational School of Health Services – Aydın, Turkey.
[2]Adiyaman University, Faculty of Medicine, Department of Obstetrics and Gynecology – Adıyaman, Turkey.
[3]Ondokuz Mayis University, Faculty of Veterinary Medicine, Department of Genetics – Samsun, Turkey.
[4]Aksaray University – Aksaray, Turkey.
[5]Adiyaman University, Faculty of Medicine, Department of Internal Medicine – Adıyaman, Turkey.
*Corresponding author: recaiaci@adu.edu.tr
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 12, 2026. Accepted on July 13, 2026.
Scientific Editor: José Maria Soares Júnior

that rs35569394 might be associated with PCOS status and explored whether genotype patterns differed according to self-reported family history and metabolic features. Any positive result was considered hypothesis-generating rather than immediately clinically actionable.

## METHODS

### Study population and design

This prospective case-control study was conducted at the Department of Obstetrics and Gynecology, Faculty of Medicine, Adiyaman University. The case group included 104 women diagnosed with PCOS according to the Rotterdam criteria, and the control group included 128 healthy women with regular menstrual cycles and no clinical, biochemical or ultrasonographic evidence of PCOS. Controls were recruited pragmatically from eligible women available during the study period. Individual matching, a fixed matching ratio, and matching calipers were not prespecified. Recruitment aimed to approximate the age and body mass index (BMI) distributions at group level, but BMI remained higher in the PCOS group. Written informed consent was obtained from all participants. The study was approved by the Adiyaman University Ethics Committee (decision date: 24 May 2022; meeting no. 5; decision no. 2022/5/19) and conducted in accordance with the Declaration of Helsinki.

### Sample size and power analysis

An a priori sample-size calculation was performed using G*Power software (version 3.1.9.4). For a chi-square test comparing the 2×3 genotype distribution between the PCOS and control groups, an effect size of Cohen's w=0.215 was assumed based on previously reported rs35569394 genotype distributions in PCOS[6]. With alpha=0.05 and 80% statistical power, the minimum required total sample size was 209. Allowing for approximately 10% incomplete or unusable data, the recruitment target was set at 230 participants. The final sample comprised 232 women (104 with PCOS and 128 controls), exceeding the planned total. Subgroup and multiplicity-adjusted analyses nevertheless had lower effective power and were interpreted cautiously.

### Eligibility and family history

PCOS was defined by the presence of at least two Rotterdam criteria: oligo-anovulation, clinical and/or biochemical hyperandrogenism, and polycystic ovarian morphology. Controls had regular 21–35-day cycles, normal ovarian morphology, and no evidence of hyperandrogenism. Although all cases met Rotterdam criteria, the four component-defined phenotypes were not recorded in a form permitting reliable retrospective A–D classification. The pooled analysis may therefore conceal phenotype-specific genetic and metabolic associations. Exclusion criteria were pregnancy, lactation, hormonal contraceptive or insulin-sensitizing drug use within the preceding 3 months, hyperprolactinemia, thyroid dysfunction, Cushing syndrome, congenital adrenal hyperplasia, chronic systemic disease, hepatic or renal failure and malignancy.

A positive family history was defined as a previous diagnosis of PCOS in the participant's mother or sister. Family-history information was obtained through structured personal interviews and was not independently verified from relatives' medical records. Family-history analyses were therefore considered exploratory and susceptible to recall and classification bias. Controls were not excluded based on a family history of PCOS, as routine verification of relatives' diagnoses was not feasible; this is acknowledged as a limitation that may attenuate the genetic contrast between groups.

### Clinical and biochemical assessment

Fasting venous blood samples were collected in the early follicular phase (cycle days 2–5). Serum was separated by centrifugation and stored at -80°C until analysis. Follicle stimulating hormone (FSH) (Elecsys FSH), LH (Elecsys LH), estradiol (Elecsys Estradiol III), prolactin (Elecsys Prolactin II), thyroid-stimulating hormone (TSH) (Elecsys TSH), total testosterone (Elecsys Testosterone II), and fasting insulin (Elecsys Insulin) were measured by electrochemiluminescence immunoassay on a Roche cobas e801 analyzer (Roche Diagnostics, Mannheim, Germany).

Manufacturer-stated analytical measuring ranges (LoD; LoQ) were as follows: FSH, 0.3–200 mIU/mL (0.3; 1.0 mIU/mL); LH, 0.3–200 mIU/mL (0.3; 1.0 mIU/mL); estradiol, 5–3,000 pg/mL (5; 25 pg/mL); prolactin, 0.094–470 ng/mL (0.094; 0.940 ng/mL); TSH, 0.005–100 µIU/mL (0.005; 0.005 µIU/mL); insulin, 0.4–1,000 µU/mL (0.4; 1.0 µU/mL); and total testosterone, 0.087–52.0 nmol/L (0.087; 0.416 nmol/L). Representative manufacturer precision [repeatability/intermediate-precision coefficient of variation (CV)] ranges were FSH, 0.8–2.1%/3.2–3.8%; LH, 0.9–2.2%/1.1–2.3%; estradiol, 1.2–8.4%/1.8–12.3%; prolactin, 1.8–3.1%/2.7–3.8%; TSH, 0.7–3.4%/1.8–11.2%; insulin, 0.8–4.3%/1.4–5.3%; and total testosterone, 1.8–11.1%/3.5–16.8%, depending on concentration.

Corresponding manufacturer reference intervals used for contextual interpretation were FSH 3.5–12.5 mIU/mL and LH 2.4–12.6 mIU/mL in the follicular phase, estradiol 20.5–62.8 pg/mL in the early follicular phase, prolactin 4.79–23.3 ng/mL in nonpregnant women, TSH 0.270–4.20 uIU/mL, fasting insulin

2.6–24.9 µU/mL, and total testosterone 0.291–1.668 nmol/L in apparently healthy women aged 20–49 years. Total testosterone values were converted from ng/mL to nmol/L using 1 ng/mL=3.467 nmol/L for reporting. Fasting glucose was measured by the hexokinase method on a Beckman Coulter AU5800 analyzer. Insulin resistance was estimated using HOMA-IR: fasting insulin (µU/mL)×fasting glucose (mg/dL)/405[14].

### Genetic analysis

Genomic DNA was isolated from ethylenediaminetetraacetic acid (EDTA)-anticoagulated peripheral blood using a DNeasy kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. DNA concentration and purity were assessed by NanoDrop spectrophotometry, and samples with adequate A260/A280 ratios were used for genotyping. The 18-bp insertion/deletion polymorphism at position -2549 in the VEGFA promoter (rs35569394) was genotyped by direct polymerase chain reaction (PCR), based on the previously published Turkish VEGFA protocol[15]. The forward primer was 5′-GCTGAGAGTGGGGCTGACTAGGTA-3′ and the reverse primer was 5′-GTTTCTGACCTGGCTATTTCCAGG-3′. PCR was performed in a final volume of 25 µL containing approximately 100 ng genomic DNA, 1×PCR buffer, 2.0 mM MgCl2, 0.2 mM of each dNTP, 20 pmol of each primer, and 1 U Taq DNA polymerase. Amplification consisted of initial denaturation at 95°C for 10 min; 35 cycles at 95°C for 45 s, 60°C for 45 s, and 72°C for 45 s; and final extension at 72°C for 10 min. PCR products were separated on a 2.5% agarose gel and visualized using a Bio-Rad ChemiDoc system. The 229-bp fragment represented the insertion allele and the 211-bp fragment represented the deletion allele; II, detection/deletion (DD), and ID genotypes were identified by the 229-bp band, the 211-bp band, or both bands, respectively. To assess reproducibility, 20% of randomly selected samples were reanalysed, with complete genotype concordance.

### Statistical analysis

Analyses were performed using International Business Machines Statistical Package for the Social Sciences (IBM SPSS Statistics), version 26.0 (IBM Corp., Armonk, NY, USA). Normality was evaluated using the Shapiro-Wilk test. Continuous variables are presented as mean +/- standard deviation and categorical variables as n (%). Group comparisons used independent-samples t-tests or chi-square tests, as appropriate. The overall 2×3 genotype distribution was evaluated before genotype-specific comparisons. Genotype and allele distributions were compared between groups, and odds ratios (ORs) with 95%CIs were calculated. Effect sizes were summarized using Cramer's V for the overall 2×3 genotype comparison and Cohen's h for the DD genotype comparison. HWE in controls was evaluated using both the asymptotic Pearson chi-square test and an exact test because results near the conventional 0.05 threshold can differ owing to the discrete genotype distribution and the approximation used by the Pearson procedure. Benjamini-Hochberg false-discovery-rate (FDR)-adjusted p-values are reported for the 11 clinical comparisons. The prespecified Bonferroni-adjusted values are retained for genotype and family-history contrasts. A single global correction across all clinical, genotype, and family-history tests was not performed because the a priori statistical plan specified domain-based corrections (FDR for clinical variables, Bonferroni for genetic and family-history contrasts) to control error rates within distinct biological families of tests. Because individual-level data were not retained for post hoc revision, a BMI-adjusted logistic regression model could not be reconstructed, which is acknowledged as a limitation. Although the final sample exceeded the a priori recruitment target, negative subgroup and multiplicity-adjusted findings may still reflect limited power.

## RESULTS

Clinical and biochemical comparisons are summarized in Table 1. After Benjamini-Hochberg FDR correction for 11 clinical comparisons, BMI, prolactin, HOMA-IR, insulin, FSH, LH and total testosterone remained statistically higher in the PCOS group than in controls (all FDR-adjusted p≤0.005). Total testosterone was 3.16 +/- 0.45 nmol/L in women with PCOS and 2.08 +/- 1.25 nmol/L in controls (raw p<0.001; FDR-adjusted p=0.002). Glucose, estradiol, TSH, and age at menarche did not differ significantly after FDR correction (FDR-adjusted p=0.165, 0.219, 0.113, and 0.627, respectively).

The overall genotype distribution differed between groups (Cramer's V=0.20; chi-square=8.87, df=2, p=0.012). The DD genotype was more frequent among PCOS patients (30.77 vs. 15.63%; Cohen's h=0.36; raw p=0.007; adjusted p=0.028; OR 2.40, 95%CI 1.28–4.51) and was the only genotype-specific comparison that remained significant after Bonferroni correction. The lower ID frequency was nominally significant but did not remain significant after correction (raw p=0.017; adjusted p=0.068). The II genotype and D-allele comparisons were not significant (adjusted p=1.000 and 0.456, respectively). The genotype distribution in controls was at the HWE boundary by the exact test and below it by the Pearson test (DD/ID/II=20/75/33; Pearson chi-square=4.34, p=0.037; exact p=0.050) (Table 2).

Family-history subgroup findings are shown in Table 3. No genotype or allele comparison remained significant after Bonferroni correction. The DD comparison between familial PCOS and controls was nominally significant (raw p=0.018) but not after adjustment (p=0.072).

**Table 1.** Clinical and laboratory findings.

| Characteristics | Patients (n=104), mean +/- SD | Controls (n=128), mean +/- SD | Raw p | FDR-adjusted p |
|---|---|---|---|---|
| BMI (kg/m²) | 27.325±3.587 | 25.23±4.271 | <0.001 | 0.002 |
| Age at menarche (year) | 12.27±0.601 | 12.36±1.72 | 0.627 | 0.627 |
| Prolactin (ng/mL) | 13.4±8.934 | 9.63±6.06 | 0.003 | 0.005 |
| HOMA-IR | 2.3±2.13 | 1.55±1.6 | 0.001 | 0.004 |
| Glucose (mg/dL) | 85.2±12.63 | 87.63±9.31 | 0.135 | 0.165 |
| Insulin (µU/mL) | 10.5±6.13 | 7.67±6.06 | 0.001 | 0.004 |
| E2 (Estradiol) (pg/mL) | 78.95±62.38 | 90.31±71.23 | 0.199 | 0.219 |
| FSH (mIU/mL) | 9.31±3.65 | 7.85±3.41 | 0.003 | 0.005 |
| LH (mIU/mL) | 10.27±9.36 | 7.6±5.35 | 0.002 | 0.004 |
| TSH (mIU/L) | 1.80±2.13 | 1.23±3.62 | 0.082 | 0.113 |
| TT (nmol/L) | 3.16±0.45 | 2.08±1.25 | <0.001 | 0.002 |

Values are presented as mean +/- standard deviation. Between-group comparisons used independent-samples t-tests. FDR-adjusted p-values were calculated by the Benjamini-Hochberg procedure for 11 clinical comparisons. Total testosterone is reported in nmol/L after conversion from ng/mL using 1 ng/mL=3.467 nmol/L. BMI: body mass index; HOMA-IR: homeostatic model assessment of insulin resistance; FSH: follicle-stimulating hormone; LH: luteinizing hormone; TSH: thyroid-stimulating hormone; E2: estradiol; TT: total testosterone; SD: standard deviation; FDR: false-discovery-rate.

**Table 2.** Genotype and allele frequencies of the Vascular Endothelial Growth Factor A-2549 I/D polymorphism.

| Genotype/allele | Patients n | Patients % | Controls n | Controls % | Chi-square | Raw p | Adjusted p | OR (95%CI) |
|---|---|---|---|---|---|---|---|---|
| DD | 32 | 30.77 | 20 | 15.63 | 7.241 | 0.007 | 0.028 | 2.40 (1.28–4.51) |
| ID | 44 | 42.31 | 75 | 58.59 | 5.650 | 0.017 | 0.068 | 0.52 (0.31–0.87) |
| II | 28 | 26.92 | 33 | 25.78 | 0.022 | 0.881 | 1.000 | 1.06 (0.59–1.90) |
| Total (genotype) | 104 | 100 | 128 | 100 | | | | |
| D allele | 108 | 51.92 | 115 | 44.92 | 2.510 | 0.114 | 0.456 | 1.35 (0.94–1.95) |
| I allele | 100 | 48.08 | 141 | 55.08 | | | | |
| Total (allele) | 208 | 100 | 256 | 100 | | | | |

Overall genotype comparison: chi-square=8.87, df=2, p=0.012. Adjusted p-values use Bonferroni correction for four genotype/allele contrasts. Control HWE: Pearson chi-square=4.34, p=0.037; exact p=0.050. OR: odds ratio; CI: confidence interval.

**Table 3.** Vascular Endothelial Growth Factor A-2549 I/D genotype and allele frequencies according to family history.

| Genotype/allele | Family history (+), n=46 | Family history (-), n=58 | Controls, n=128 | P1 raw (adjusted) | OR (95%CI) | P2 raw (adjusted) | OR (95%CI) |
|---|---|---|---|---|---|---|---|
| DD | 15 | 17 | 20 | 0.831 (1.000) | 1.17 (0.51–2.67) | 0.018 (0.072) | 2.61 (1.21–5.66) |
| ID | 21 | 23 | 75 | 0.556 (1.000) | 1.28 (0.59–2.78) | 0.167 (0.668) | 0.59 (0.30–1.16) |
| II | 10 | 18 | 33 | 0.374 (1.000) | 0.62 (0.25–1.50) | 0.692 (1.000) | 0.80 (0.36–1.78) |
| Total (genotype) | 46 | 58 | 128 | | | | |
| D allele | 51 | 57 | 115 | | | | |
| I allele | 41 | 59 | 141 | 0.403 (1.000) | 1.29 (0.75–2.22) | 0.090 (0.360) | 1.53 (0.95–2.46) |
| Total (allele) | 108 | 100 | 256 | | | | |

P1 compares PCOS patients with and without self-reported family history; P2 compares PCOS patients with positive family history and controls. Values in parentheses are Bonferroni-adjusted p-values across four contrasts. OR: odds ratio; CI: confidence interval.

## DISCUSSION

In this Turkish case-control cohort, the DD genotype of VEGFA -2549 I/D was more frequent among women with PCOS than among controls, whereas the allele-level comparison was not significant. The overall genotype distribution also differed between groups. These findings support a possible genotype-specific association but do not establish causality. Because rs35569394 is located in the VEGFA promoter, it has been

hypothesized to modify transcriptional regulation; however, VEGFA expression, circulating protein concentration, ovarian tissue activity, and promoter function were not measured in this study. The mechanistic interpretation therefore remains biologically plausible but unconfirmed.

The clinical profile of the PCOS group was characterized by higher BMI, insulin, HOMA-IR, LH, and total testosterone, consistent with the endocrine-metabolic heterogeneity of PCOS[16-20]. Hyperinsulinemia may intensify ovarian androgen production and reduce hepatic sex hormone-binding globulin (SHBG), while adiposity can further modify gonadotropin and metabolic patterns. Total testosterone was higher in PCOS than in controls (3.16 +/- 0.45 vs. 2.08 +/- 1.25 nmol/L; FDR-adjusted p=0.002), supporting a group-level hyperandrogenic difference. Although the mean total testosterone in the control group slightly exceeded the manufacturer's general reference interval, it was significantly lower than in the PCOS group and remained within the expected physiological spectrum for reproductive-aged women in our laboratory setting. FSH was also modestly higher in cases. Because increased FSH is not a defining PCOS feature, this observation may reflect phenotype composition, cycle timing, assay characteristics, adiposity, or sampling variation[21]. The absence of free testosterone, SHBG, repeated measurements, and phenotype-specific hormonal profiles limits more detailed endocrine interpretation.

VEGFA is biologically relevant to PCOS because follicular development, ovulation, and corpus luteum formation depend on tightly regulated angiogenesis. Excessive or mistimed VEGFA signalling could contribute to increased stromal vascularity, vascular permeability, altered follicular maturation, and an exaggerated ovarian response. Contemporary granulosa-cell and ovarian-stimulation studies show that angiogenesis-related gene expression and VEGF-mediated vascular responses vary according to cellular and treatment context[22,23]. These observations support the ovarian relevance of the pathway but do not demonstrate that rs35569394 produces the same effects in PCOS. Associations involving circulating VEGF or other VEGFA variants have also varied among populations[6-12], suggesting that ancestry, linkage disequilibrium, clinical phenotype, adiposity, and analytical methods may influence observed genetic effects.

The control-group HWE deviation remains an important methodological limitation. The exact test was borderline (p=0.050), whereas the Pearson test was below 0.05 (p=0.037). Complete concordance in the 20% repeat-genotyping subset reduces concern about random technical error but does not exclude systematic misclassification, population stratification, cryptic relatedness, or selection effects. The control genotype distribution may therefore not represent the underlying population frequencies, and the observed odds ratio should be interpreted cautiously.

The D-allele frequency was 51.9% among cases and 44.9% among controls in the present cohort. By comparison, Fernandes et al. reported a D-allele frequency of 0.64 in both Brazilian cases and controls and found no association with PCOS[7]. Their control group also departed from HWE. Differences in allele distribution and genotype structure between the Brazilian and Turkish cohorts illustrate the importance of ancestry, recruitment, and sampling framework in interpreting rs35569394. The present DD association, despite a non-significant allele comparison, may be compatible with a recessive or genotype-specific pattern; alternatively, it may result from limited power, control selection, residual confounding, or HWE deviation. Accordingly, replication should test prespecified dominant, recessive, and additive genetic models rather than focusing only on isolated genotype contrasts.

An important interpretive issue is residual confounding by adiposity. Although the groups were broadly similar in age, BMI was significantly higher in the PCOS group. BMI may influence both PCOS status and metabolic or angiogenic phenotypes; therefore, the observed DD association cannot be interpreted as independent of adiposity because an individual-level covariate-adjusted model was not available for revision. Future studies should use BMI-matched recruitment or multivariable logistic regression to estimate the independent genetic effect.

Family history was more common among women with PCOS, but the DD comparison in the familial subgroup did not remain significant after multiplicity correction. Familial aggregation is consistent with the complex genetic architecture of PCOS, yet self-reported history is an imprecise proxy for inherited susceptibility because relatives' diagnoses were not verified and shared environmental factors cannot be separated from genetic effects. The family-history analysis should therefore be viewed as exploratory rather than evidence that rs35569394 explains familial PCOS.

Taken together, the results identify rs35569394 as a candidate marker warranting further investigation in Turkish women rather than as a diagnostic or predictive test. A clinically useful marker would require reproducible effects across independent cohorts, stability after adjustment for adiposity and ancestry, consistency across Rotterdam phenotypes, and evidence linking genotype to VEGFA expression or ovarian function. The present study provides an initial population-specific signal but does not meet those requirements for clinical translation.

### Limitations

The main limitations were the control-group HWE deviation, incomplete BMI matching without an available covariate-adjusted model, domain-based rather than global multiplicity correction, and limited power for subgroup and multiplicity-adjusted analyses. These factors increase the possibility of genotyping, selection, confounding, or type I/II error despite repeat-genotyping concordance.

Rotterdam A–D phenotypes could not be reconstructed reliably, family history was self-reported, and no VEGFA expression or functional assays were performed. The pooled association should therefore be considered preliminary and not extrapolated to a particular phenotype or to clinical utility[24,25]. Larger, well-matched studies with prespecified quality control, ancestry assessment, phenotype stratification, unified multiplicity correction, and functional validation are needed.

## CONCLUSION

In this Turkish case-control sample, the VEGFA -2549 I/D DD genotype was more frequent among women with PCOS than among controls. This study shows an observed association; it does not establish causality, functional effects, phenotype-specific relevance, or clinical utility. Control-group HWE deviation, residual confounding from incomplete matching, limited power, non-global multiplicity correction, absence of phenotype stratification, and lack of functional validation make the finding preliminary. Independent replication in larger, well-matched cohorts with prespecified genotyping quality control, ancestry assessment, covariate analysis, phenotype stratification, and functional studies is required.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

This study was approved by the Ethics Committee of Adiyaman University Faculty of Medicine (Decision Number: 2022/5/19; Decision Date: 24/05/2022). The study was conducted in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrollment.

## AUTHORS' CONTRIBUTIONS

**RA:** Conceptualization, Project administration, Resources, Validation, Writing - original draft. **MCN:** Conceptualization, Project administration, Resources, Validation, Writing - original draft. **SY:** Data curation, Formal Analysis, Resources. **HR:** Conceptualization, Validation, Writing - review & editing. **EN:** Formal Analysis, Software, Visualization.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Pasquali R, Gambineri A. New perspectives on the definition and management of polycystic ovary syndrome. J Endocrinol Invest. 2018;41:1123-35.
2. Mahabamunuge J, Sekula NM, Lepore C, Kudrimoti M, Upadhyay A, Alshowaikh K, et al. The molecular basis of polycystic ovary syndrome and its cardiometabolic correlates. Biomedicines. 2025;13:709.
3. Szkodziak P, Szkodziak F, Trzeciak K, Wozniak S, Mlynarczyk M, Paszkowski T. Insulin resistance in polycystic ovary syndrome phenotypes and the vicious cycle model. Sci Rep. 2025;15:42649.
4. Bozdag G, Mumusoglu S, Zengin D, Karabulut E, Yildiz BO. The prevalence and phenotypic features of polycystic ovary syndrome: a systematic review and meta-analysis. Hum Reprod. 2016;31:2841-55.
5. Gleicher N, Darmon S, Patrizio P, Barad DH. Reconsidering the polycystic ovary syndrome. Biomedicines. 2022;10:1505.
6. Huang L, Wang L. Association between VEGF gene polymorphisms and polycystic ovary syndrome risk. Biosci Rep. 2020;40:BSR20191691.
7. Fernandes SS, Marqui ABT, Teles DRF, Resende EAMR, Lima MFP, Gomes MKO, et al. VEGF gene rs35569394 polymorphism in patients with polycystic ovary syndrome. Rev Assoc Med Bras. 2020;66:1396-401.
8. Zhao J, Li D, Tang H, Tang L. Association of vascular endothelial growth factor polymorphisms with polycystic ovarian syndrome risk: a meta-analysis. Reprod Biol Endocrinol. 2020;18:18.
9. Ben Salem A, Megdich F, Kacem O, Souayeh M, Hachani Ben Ali F, Hizem S, et al. VEGFA gene variation in polycystic ovary syndrome in a Tunisian women population. BMC Genomics. 2016;17(Suppl 9):748.
10. Almawi WY, Gammoh E, Malalla ZH, Al-Madhi SA. VEGFA variants and changes in VEGF levels underscore the contribution of VEGF to polycystic ovary syndrome. PLoS One. 2016;11:e0165636.
11. Skrzynska K, Zachurzok A, Pietrusik A, Jakubowska-Kowal K, Gawlik-Starzyk A. Visfatin and VEGF levels are not increased in adolescent girls with polycystic ovary syndrome. Front Endocrinol. 2024;15:1488249.
12. Vicente ALSA, Marqui ABT, Gomes MKO, Assunção-Luiz AV, Balarin MAS, Tanaka SCSV, et al. Polymorphisms rs2010963 and rs833061 of the VEGF gene in polycystic ovary syndrome. Rev Assoc Med Bras (1992). 2022;68(6):785-91. https://doi.org/10.1590/1806-9282.20211345
13. Almotairi R, Mir R, Almasoudi KS, Husain E, Mtiraoui N. Evaluating the impact of rs4025935, rs71748309, rs699947,

and rs4646994 genetic determinants on PCOS predisposition. Life. 2025;15:558.

14. Bonora E, Targher G, Alberiche M, Bonadonna RC, Saggiani F, Zenere MB, et al. Homeostasis model assessment closely mirrors the glucose clamp technique. Diabetes Care. 2000;23:57-63.
15. Yigit S, Nursal AF, Celik A, Aci R, Askeroglu E. Lack of association between the -2549 insertion/deletion variant of vascular endothelial growth factor and coronary artery disease in the Turkish population. Rev Assoc Med Bras. 2024;70:e20240333.
16. Lewandowski KC, Skowronska-Jozwiak E, Lukasiak K, Galuszko K, Dukowicz A, Cedro M, et al. How much insulin resistance in polycystic ovary syndrome? Comparison of HOMA-IR and insulin resistance (Belfiore) index models. Arch Med Sci. 2019;15:613-8.
17. Bahadur A, Verma N, Mundhra R, Chawla L, Ajmani M, Sri MS, et al. Correlation of HOMA-IR, anti-Mullerian hormone and BMI in PCOS characterization. Cureus. 2021;13:e15972.
18. Unfer V, Kandaraki E, Pkhaladze L, Roseff S, Vazquez-Levin MH, Lagana AS, et al. When one size does not fit all: reconsidering PCOS etiology and clinical subgroups. Endocr Metab Sci. 2024;14:100159.
19. Elsayed AM, Al-Kaabi LS, Al-Abdulla NM, Al-Kuwari MS, Al-Mulla AA, Al-Shamari RS, et al. Clinical phenotypes of PCOS: a cross-sectional study. Reprod Sci. 2023;30:3261-72.
20. Toulis KA, Goulis DG, Mintziori G, Kintiraki E, Eukarpidis E, Mouratoglou SA, et al. Meta-analysis of cardiovascular disease risk markers in women with PCOS. Hum Reprod Update. 2011;17:741-60.
21. Saadia Z. Follicle stimulating hormone (LH:FSH) ratio in polycystic ovary syndrome: obese vs non-obese women. Med Arch. 2020;74:289.
22. Cavalcanti GS, Carvalho KC, Monteleone PAA, Chedraui P, Cipolla-Neto J, Ueno J, et al. Melatonin modulates granulosa cell angiogenesis differently in female- and male-factor infertility. Gynecol Endocrinol. 2026;42:2606440.
23. Beltrame AL, Serafini P, Motta EL, Soares Júnior JM, Baracat EC. The effects of bromocriptine on VEGF, kidney function and ovarian hyperstimulation syndrome in in vitro fertilization patients: a pilot study. Gynecol Endocrinol. 2013;29(3):201-4. https://doi.org/10.3109/09513590.2012.736554
24. Zuchelo LTS, Alves MS, Baracat EC, Sorpreso ICE, Soares Junior JM. Menstrual pattern in polycystic ovary syndrome and hypothalamic-pituitary-ovarian axis immaturity in adolescents: a systematic review and meta-analysis. Gynecol Endocrinol. 2024;40:2360077.
25. Baracat EC, Baracat MCP, Soares Junior JM. Are there new insights for the definition of PCOS? Gynecol Endocrinol. 2022;38:703-4.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260845

# Weight loss–independent associations of body composition changes with device-derived metabolic age during Mediterranean diet follow-up

Ahmet Emre Hatır[1]*, Mustafa Onmaz[1], Reyhan Vildan Acar[2]

## SUMMARY

**OBJECTIVE:** The aim of this study was to investigate whether changes in body composition over 6 months of Mediterranean diet follow-up were associated with corresponding changes in device-derived metabolic age after adjustment for weight change.

**METHODS:** This retrospective observational study analyzed body composition records of 414 adults who completed routine 6-month follow-up after Mediterranean diet counseling between June 2024 and June 2025. Metabolic age and body composition parameters were assessed using the Tanita MC-580 body composition analyzer at baseline and after 6 months. Hierarchical multiple linear regression analysis was used to evaluate the associations between changes in body composition parameters and metabolic age change after adjustment for weight change, age, sex, and baseline metabolic age.

**RESULTS:** Tanita-derived metabolic age decreased by 0.94 years ($p<0.001$). Changes in muscle–fat ratio ($\beta$=-0.585), visceral fat ($\beta$=-0.556), and protein mass ($\beta$=-0.339) were significantly associated with metabolic age change after adjustment, whereas weight-adjusted waist index was not significantly associated with metabolic age change in the adjusted model. Inclusion of body composition variables increased the explained variance of the model from 20.4 to 45.7% (adjusted $R^2$ from 0.196 to 0.446).

**CONCLUSION:** Favorable changes in Tanita-derived metabolic age were observed during follow-up. The associated body composition changes may partly reflect shared components of the device algorithm because metabolic age is a composite surrogate rather than a direct measure of biological aging. Given the retrospective uncontrolled design, causal effects of the Mediterranean diet cannot be established.

**KEYWORDS:** Body composition. Electric impedance. Mediterranean diet. Visceral fat.

## INTRODUCTION

Metabolic age is often used as a simple indicator of overall physical and metabolic condition. It is estimated by comparing an individual's basal metabolic rate (BMR) with the expected BMR for people of the same chronological age. In general, a lower metabolic age is thought to be associated with a more favorable cardiometabolic profile. However, this measure should be interpreted cautiously, as it is a device-derived estimate shaped by body composition and related physiological features rather than a direct biological marker. Factors such as genetics, muscle and fat mass, physical activity, and dietary habits may all influence metabolic age[1,2]. Among dietary approaches, the Mediterranean diet has received particular attention because of its associations with improvements in body weight, lipid profile, blood pressure, and insulin sensitivity[3]. Accordingly, changes in this device-derived estimate during dietary follow-up may coincide with both weight change and changes in body composition.

Body composition, especially the balance between fat mass and lean tissue, plays an important role in metabolic health. Visceral fat is closely linked to insulin resistance, dyslipidemia, and higher cardiovascular risk, even in individuals without marked excess body weight. In contrast, a greater proportion of lean mass is generally associated with better metabolic regulation and physical performance. Evidence from recent studies suggests that dietary interventions can improve metabolic outcomes by reducing harmful fat depots, particularly visceral fat, even when total weight loss is relatively modest[3,4]. Beyond these major body compartments, body water compartments are also measured by bioelectrical impedance analysis (BIA) and may vary with tissue composition and fluid distribution[5,6].

Although metabolic age is increasingly used as a simple device-derived indicator in body composition assessment, its clinical interpretation after dietary follow-up remains insufficiently defined. In routine practice, dietary response is often evaluated primarily through body weight, body mass index (BMI), and

[1]Bandırma Onyedi Eylul University, Department of Family Medicine – Balıkesir, Türkiye.
[2]Hasan Katırcı Healthy Living Center – Konya, Türkiye.
*Corresponding author: ahatir@bandirma.edu.tr
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on May 13, 2026. Accepted on July 13, 2026.
Scientific Editor: José Maria Soares Júnior

waist circumference, whereas these parameters may not fully capture qualitative changes in body composition. Importantly, because device-derived metabolic age is calculated using BMR and body-composition-related inputs, associations between this estimate and BIA-derived variables may partly reflect shared algorithmic components. Nevertheless, examining these adjusted associations may help clarify how the device-derived estimate behaves during routine follow-up. Therefore, the present study aimed to examine the associations between changes in body composition parameters and changes in Tanita-derived metabolic age during 6-month Mediterranean diet follow-up after accounting for weight change and baseline covariates.

## METHODS

This retrospective observational study was based on the analysis of body composition records from individuals who participated in a Mediterranean diet follow-up between June 2024 and June 2025. Measurements obtained at baseline (month 0) and at 6 months were analyzed. Participants were not prospectively randomized or allocated by the investigators for research purposes; rather, routinely collected records of individuals who had attended a dietitian-led Mediterranean diet follow-up program were analyzed. During the study period, all available records of individuals attending the Mediterranean diet follow-up were screened sequentially. Of 465 baseline records screened, 51 were excluded because of missing 6-month measurements, missing or erroneous data, or predefined exclusion criteria. The final complete-case cohort included 414 participants, representing 89.0% of the screened baseline records. Because this was a retrospective study based on routinely collected records, no a priori sample size calculation was performed. Instead, all eligible participants with complete baseline and 6-month measurements during the predefined study period were included. Participants followed a nutritional program based on Mediterranean diet principles under the guidance of a registered dietitian. Dietary counseling emphasized increased consumption of vegetables, fruits, legumes, whole grains, olive oil, nuts, and fish, together with reduced intake of refined carbohydrates, processed foods, sugar-sweetened products, red and processed meats, and saturated fat-rich foods. Adherence was assessed during routine dietitian-led follow-up visits based on the dietary information documented in the clinical records; no validated Mediterranean diet adherence scale, dietary recall, food diary, food-frequency questionnaire, or objective nutritional biomarker was used. Changes in participants' physical activity habits and medication use were not systematically monitored. Chronic disease status and laboratory test results were not incorporated into the analyses. Participants were eligible if they were aged 18 years or older, had undergone multi-frequency BIA at both time points, had documented follow-up for 6 months, and met the pre-measurement fasting requirement of 8–12 h. Individuals were excluded if they had a history of acute infection, malignancy, renal or hepatic failure, or other systemic diseases; were pregnant or breastfeeding; had pacemakers or metallic implants; had severe edema or fluid imbalance; had incomplete or erroneous follow-up records; or were younger than 18 years of age.

Body composition was measured using a segmental bioelectrical impedance analyzer (Tanita MC-580; Tanita Corporation, Tokyo, Japan), which employs an eight-electrode hand-to-foot system and dual-frequency BIA (6.25 and 50 kHz)[7-9]. All measurements were taken in the morning after fasting for 8 to 12 h. Participants were asked to avoid vigorous physical activity in the past 24 h, empty their bladder, and remove metallic accessories prior to the test. Room temperature was maintained between 22 and 24°C. Age, height, and sex were entered into the device before measurement. The analyzer was calibrated according to the manufacturer's instructions, and all measurements were performed by the same operator. BMI was computed as weight in kilograms divided by height in meters squared. Weight-adjusted waist index (WWI) was computed by dividing waist circumference in centimeters by the square root of body weight in kilograms. Muscle-to-fat ratio (MFR) was defined as skeletal muscle mass divided by fat mass. Metabolic age is a proprietary composite surrogate estimate derived by the device from BMR and body-composition-related parameters, including fat-free mass and body fluid distribution. It should therefore be interpreted as a device-derived output rather than as a direct measure of biological or epigenetic aging processes[9].

The study was conducted in accordance with the Declaration of Helsinki and the national ethical standards of Turkey, following ethical approval from the Ethics Committee for Non-Interventional Clinical Research at KTO Karatay University, Faculty of Medicine (Approval No: 2026/032, protocol number 132293, date: 26 February 2026).

### Statistical analysis

Data were analyzed using International Business Machines Statistical Package for the Social Sciences Statistics for Windows, version 27.0 (IBM Corp., Armonk, NY, USA). Figure 1 was generated using Python (version 3.10). Even though normality tests were statistically significant, skewness and kurtosis were within acceptable limits; parametric analyses were therefore considered robust. Continuous variables were presented as mean±standard deviation. The primary outcome was metabolic age change, calculated as the 6-month value minus the baseline value (Δ).

Paired comparisons were performed using the paired samples t-test. Hierarchical multiple linear regression analysis was used with metabolic age change as the dependent variable. For each regression model, unstandardized coefficients (B), standard errors (SEs), standardized coefficients (β), 95%CIs, p-values, and variance inflation factor (VIF) values were reported. Multicollinearity was assessed using VIF values. In the base model, weight change, age, sex, and baseline metabolic age were included. In Model 2, changes in protein mass, MFR, WWI, and visceral fat were added to evaluate their associations beyond weight change and baseline covariates. Candidate variables were selected based on clinical and physiological relevance, avoiding overfitting and multicollinearity[10]. A sensitivity power analysis was conducted using G*Power version 3.1 to determine the minimum incremental effect size detectable with the available sample in the hierarchical multiple regression analysis. With 414 participants, an alpha level of 0.05, four tested predictors, eight total predictors, and 80% statistical power, the minimum detectable incremental effect size was $f^2$=0.029, corresponding to a small effect. The significance level was set at p<0.05.

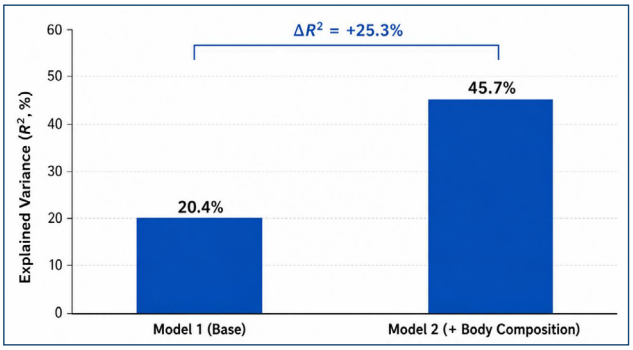



**Figure 1.** Incremental explained variance across hierarchical regression models. Explained variance across hierarchical regression models for change in device-derived metabolic age. Model 1 included baseline covariates and weight change. Model 2 additionally incorporated body composition variables, increasing the explained variance from $R^2$=0.204 to $R^2$=0.457 ($\Delta R^2$=0.253).

## RESULTS

A total of 414 individuals were included. Participants had a mean age of 40.78±10.86 years (range 19–70). The study population comprised 342 women (82.6%) and 72 men (17.4%).

Significant changes were observed in metabolic age and most body composition parameters during follow-up. Metabolic age decreased significantly (-0.94 years, p<0.001). Substantial reductions were noted in body weight (-2.60 kg, p<0.001) and BMI (-1.02 kg/m$^2$, p<0.001). Waist circumference, fat mass, visceral fat, and muscle mass indicators also changed significantly during follow-up (all p<0.001). MFR improved significantly, reflecting a more favorable lean-to-fat balance (p<0.001). WWI decreased significantly from 10.90±0.79 to 10.76±0.80 (Δ=-0.14, p<0.001), indicating an improvement in weight-adjusted central adiposity. Total body water showed a modest but significant decrease (p<0.001) (Table 1).

**Table 1.** Changes in anthropometric and body composition parameters during Mediterranean diet follow-up.

| Variable | Month 0 Mean±SD | Month 6 Mean±SD | Mean Change (Δ) | p |
|---|---|---|---|---|
| Metabolic age (years) | 47.81±11.69 | 46.86±11.93 | -0.94 | <0.001 |
| Weight (kg) | 83.25±17.21 | 80.65±16.65 | -2.60 | <0.001 |
| BMI (kg/m$^2$) | 32.49±6.08 | 31.46±5.94 | -1.02 | <0.001 |
| Waist circumference (cm) | 99.52±17.04 | 96.69±16.72 | -2.83 | <0.001 |
| Fat mass (kg) | 31.29±11.07 | 29.39±10.92 | -1.90 | <0.001 |
| Muscle mass (kg) | 49.26±7.80 | 48.59±7.62 | -0.67 | <0.001 |
| Skeletal muscle mass (kg) | 29.42±4.68 | 29.02±4.57 | -0.40 | <0.001 |
| Protein mass (kg) | 12.01±2.93 | 11.95±2.80 | -0.06 | 0.267 |
| Mineral mass (kg) | 3.65±0.63 | 3.60±0.61 | -0.05 | <0.001 |
| Muscle–fat ratio | 1.02±0.30 | 1.09±0.35 | +0.07 | <0.001 |
| Weight-adjusted waist index | 10.90±0.79 | 10.76±0.80 | -0.14 | <0.001 |
| Visceral fat | 8.64±4.13 | 8.02±4.02 | -0.62 | <0.001 |
| Total body water (kg) | 36.31±5.64 | 35.72±5.57 | -0.59 | <0.001 |

Changes in anthropometric and body composition parameters during 6-month Mediterranean diet follow-up. Values are presented as mean±standard deviation (SD). Mean change (Δ) represents the difference between Month 6 and baseline measurements. p-values were obtained using paired statistical tests. BMI: body mass index.

Hierarchical multiple linear regression analysis was performed to evaluate factors associated with metabolic age change after accounting for weight change and baseline covariates. In the base model (Model 1), which included weight change, age, sex, and baseline metabolic age, the model explained 20.4% of the variance in metabolic age change (adjusted $R^2$=0.196, p<0.001). Weight change (β=0.459, p<0.001) and baseline metabolic age (β=0.269, p=0.032) showed statistically significant associations with metabolic age change. When body composition markers were added in Model 2, the explained variance increased from 20.4 to 45.7% (adjusted $R^2$=0.446; $\Delta R^2$=0.253, p<0.001). Changes in muscle–fat ratio (β=-0.585, p<0.001), visceral fat (β=-0.556, p<0.001), and protein mass (β=-0.339, p<0.001) remained significantly associated with metabolic age change after adjustment. ΔWWI was not statistically significant (β=-0.135, p=0.124) (Table 2). The incremental explained variance across hierarchical regression models is illustrated in Figure 1.

## DISCUSSION

During 6 months of Mediterranean diet follow-up, changes in MFR, visceral fat, and protein mass remained significantly associated with changes in Tanita-derived metabolic age after adjustment for weight change and baseline covariates. These findings are consistent with previous studies reporting favorable changes in body composition, fat distribution, and metabolic risk profiles during Mediterranean diet interventions[4,11]. Broader literature shows that adherence to a Mediterranean diet is associated with reduced central adiposity, improved metabolic syndrome markers, and lower cardiovascular risk[11-15], and that combined lifestyle interventions including the Mediterranean diet improve visceral adiposity while preserving lean mass[4]. Importantly, the associations between BIA-derived body composition variables and changes in Tanita-derived metabolic age may partly reflect shared measurement and algorithmic components. The regression findings should therefore be interpreted as adjusted associations or correlates of change in the device-derived estimate, rather than as evidence that these variables independently determine metabolic or biological aging.

MFR change showed the largest standardized association with metabolic age change among the body composition variables included in Model 2. This is physiologically plausible because skeletal muscle plays a central role in glucose utilization, insulin sensitivity, and resting energy expenditure, whereas excess adiposity is associated with chronic low-grade inflammation[16,17]. Nevertheless, because MFR and metabolic age were both derived from the same BIA assessment, this association should be interpreted cautiously and may partly reflect shared algorithmic components. The adjusted association between visceral fat change and metabolic age change is another noteworthy finding. Visceral adiposity is associated with insulin resistance, dyslipidemia, fatty liver, and inflammation[17,18]. A previous intervention study reported reductions in visceral adiposity during Mediterranean diet follow-up, including in the presence of moderate weight loss[4]. Our findings that visceral fat reduction and improved muscle–fat ratio were associated with changes in device-derived metabolic age are consistent with observations from

**Table 2.** Hierarchical multiple linear regression analysis of factors associated with Δmetabolic age.

| | Variables | B | SE | β | 95%CI (LL–UL) | p | VIF | Adj. $R^2$ |
|---|---|---|---|---|---|---|---|---|
| Model 1 Base model (weight change controlled) | ΔWeight | 0.321 | 0.032 | 0.459 | 0.259–0.384 | <0.001 | 1.060 | 0.196 |
| | Age | -0.029 | 0.021 | -0.175 | -0.069–0.012 | 0.161 | 7.964 | |
| | Sex | 0.120 | 0.212 | 0.025 | -0.297–0.537 | 0.572 | 1.033 | |
| | Baseline metabolic age | 0.041 | 0.019 | 0.269 | 0.004–0.079 | 0.032 | 8.063 | |
| Model 2 (+body composition) | ΔProtein | -0.522 | 0.099 | -0.339 | -0.717 to -0.327 | <0.001 | 3.096 | 0.446 |
| | ΔMuscle–fat ratio | -7.544 | 0.959 | -0.585 | -9.429 to -5.659 | <0.001 | 4.123 | |
| | ΔWeight-adjusted waist index | -0.957 | 0.621 | -0.135 | -2.178 to 0.264 | 0.124 | 5.729 | |
| | ΔVisceral fat | -0.935 | 0.121 | -0.556 | -1.173 to -0.698 | <0.001 | 3.843 | |

B represents the unstandardized regression coefficient; β represents the standardized coefficient. SE: standard error; CI: confidence interval; VIF: variance inflation factor. Model 1 included weight change, age, sex, and baseline metabolic age. Model 2 additionally included changes in protein mass, muscle–fat ratio, weight-adjusted waist index, and visceral fat. LL: lower limit; UL: upper limit.

other metabolic populations. For example, Soares-Jr et al. reported differences in type 2 diabetes risk across polycystic ovary syndrome phenotypes despite similar BMI values, highlighting that metabolic risk may vary across phenotypic subgroups[19]. This supports the broader concept that metabolic risk assessment should consider phenotype-related characteristics beyond body weight alone.

Regarding WWI, this analysis showed a significant decrease during follow-up, suggesting an improvement in weight-adjusted central adiposity. WWI has been proposed as an anthropometric index that reflects central adiposity adjusted for body weight and has been associated with cardiometabolic outcomes in previous studies[20]. However, in the adjusted regression model, ΔWWI was not significantly associated with metabolic age change after accounting for weight change and other body composition parameters. Therefore, although WWI improved at the group level, its association with metabolic age change was attenuated when visceral fat, MFR, and protein mass were considered together.

Although mean protein mass did not change significantly at the group level, interindividual changes in BIA-derived protein mass were associated with metabolic age change in Model 2. Because both variables were estimated using the same BIA device, shared measurement or algorithmic components may partly account for this association. More broadly, Mediterranean dietary patterns, which emphasize unsaturated fats and plant-based protein sources, have been associated with favorable metabolic and body composition outcomes[15,21].

Several limitations should be acknowledged. The retrospective observational design and absence of a control group preclude causal inference; therefore, the observed changes in metabolic age and body composition cannot be attributed definitively to the Mediterranean diet. The observed changes may also have been influenced by regression to the mean, concurrent lifestyle changes, natural temporal variation, or other unmeasured factors. Dietary adherence was not assessed using a validated Mediterranean diet adherence scale, dietary recall, food-frequency questionnaire, or objective nutritional biomarker. In addition, changes in physical activity and medication use were not systematically monitored, and chronic disease status was not incorporated into the analyses. These unmeasured or incompletely measured factors may have confounded the observed associations between changes in body composition and metabolic age. Furthermore, the shared device-derived nature of the outcome and several explanatory variables creates a risk of conceptual and statistical circularity, which limits the interpretation of the observed associations. Biochemical markers such as insulin resistance indices, lipid parameters, inflammatory markers, or hormonal variables were not available, limiting mechanistic interpretation. Several body composition variables were physiologically interrelated, which may have increased multicollinearity. Moreover, because women constituted 82.6% of the sample, the findings may have limited generalizability to men and to populations with a more balanced sex distribution. Finally, the device-specific nature of metabolic age may also limit generalizability across different BIA devices or biological aging measures. Future prospective controlled studies should use validated or objective measures of Mediterranean diet adherence, systematically monitor physical activity and medication changes, and incorporate biochemical or biological aging markers that are independent of the body composition device.

In this retrospective cohort, favorable changes in body composition and Tanita-derived metabolic age were observed during 6-month Mediterranean diet follow-up. Changes in MFR, visceral fat, and protein mass were associated with metabolic age change after adjustment for weight change and baseline covariates, whereas WWI was not significantly associated in the adjusted model. These findings should be interpreted as correlates of a composite device-derived surrogate estimate rather than as evidence that body composition variables independently determine biological aging. Given the uncontrolled design, potential unmeasured confounding, and overlap with the device algorithm, causal conclusions regarding the Mediterranean diet cannot be drawn.

## AUTHORS' CONTRIBUTIONS

**AEH:** Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **MO:** Conceptualization, Investigation, Methodology, Supervision, Validation, Writing – review & editing. **RVA:** Data curation, Formal Analysis, Resources, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Kaziuk MB, Kosiba W, Kuzniewski M. High metabolic age and excessive adipose tissue as a storage location for toxins, and their influence on the excretory function of a liver and bile ducts in patients with a transplanted allogenic kidney. Transplantation. 2017;101:S53.
2. Ramírez-Gallegos I, Marina-Arroyo M, López-González ÁA, Vallejos D, Martínez-Almoyna-Rifá E, Tárraga López PJ, et al. Associations between metabolic age, sociodemographic variables, and lifestyle factors in Spanish workers. Nutrients. 2024;16:4207.
3. Scaglione S, Chiara T, Daidone M, Tuttolomondo A. Effects of the Mediterranean diet on the components of metabolic syndrome concerning the cardiometabolic risk. Nutrients. 2025;17:358.
4. Konieczna J, Ruiz-Canela M, Galmes-Panades AM, Abete I, Babio N, Fiol M, et al. An energy-reduced Mediterranean diet, physical activity, and body composition: an interim subgroup analysis of the PREDIMED-Plus randomized clinical trial. JAMA Netw Open. 2023;6:e2337994.
5. Ge YZ, Ruan GT, Zhang Q, Dong WJ, Zhang X, Song MM, et al. Extracellular water to total body water ratio predicts survival in cancer patients with sarcopenia: a multicenter cohort study. Nutr Metab (Lond). 2022;19:34.
6. Harimawan AIW, Prabandari AASM, Wihandani DM, Jawi IM, Weta IW, Senapathi TGA, et al. Association between phase angle and ECW/TBW ratio with body composition in individuals with central obesity: a cross-sectional study. Front Nutr. 2025;12:1638075.
7. Kyle UG, Bosaeus I, Lorenzo AD, Deurenberg P, Elia M, Gómez JM, et al. Bioelectrical impedance analysis—part I: review of principles and methods. Clin Nutr. 2004;23:1226-43.
8. Kyle UG, Bosaeus I, Lorenzo AD, Deurenberg P, Elia M, Gómez JM, et al. Bioelectrical impedance analysis—part II: utilization in clinical practice. Clin Nutr. 2004;23:1430-53.
9. Tanita Corporation. MC-580 body composition analyzer: instruction manual [Internet]. Tokyo (JP): Tanita Corporation; 2018. [cited on 2026 Mar 15]. Available from: https://cdn.shopify.com/s/files/1/0835/0462/7999/files/MC-580_Manual.pdf?v=1736269266
10. Babyak MA. What you see may not be what you get: a brief, nontechnical introduction to overfitting in regression-type models. Psychosom Med. 2004;66:411-21.
11. Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, Fitó M, Chiva-Blanch G, et al. Effect of a high-fat Mediterranean diet on bodyweight and waist circumference: a prespecified secondary outcomes analysis of the PREDIMED randomised controlled trial. Lancet Diabetes Endocrinol. 2019;7:e6-17.
12. Bakaloudi DR, Chrysoula L, Kotzakioulafi E, Theodoridis X, Chourdakis M. Impact of the level of adherence to Mediterranean diet on the parameters of metabolic syndrome: a systematic review and meta-analysis of observational studies. Nutrients. 2021;13:1514.
13. Zelicha H, Klöting N, Kaplan A, Yaskolka Meir A, Rinott E, Tsaban G, et al. The effect of high-polyphenol Mediterranean diet on visceral adiposity: the DIRECT PLUS randomized controlled trial. BMC Med. 2022;20:327.
14. Sebastian SA, Padda I, Johal G. Long-term impact of Mediterranean diet on cardiovascular disease prevention: a systematic review and meta-analysis of randomized controlled trials. Curr Probl Cardiol. 2024;49:102509.
15. Papadaki A, Nolen-Doerr E, Mantzoros CS. The effect of the Mediterranean diet on metabolic health: a systematic review and meta-analysis of controlled trials in adults. Nutrients. 2020;12:3342.
16. Srikanthan P, Karlamangla AS. Relative muscle mass is inversely associated with insulin resistance and prediabetes: findings from the Third National Health and Nutrition Examination Survey. J Clin Endocrinol Metab. 2011;96:2898-903.
17. Després JP. Body fat distribution and risk of cardiovascular disease: an update. Circulation. 2012;126:1301-13.
18. Neeland IJ, Ross R, Després JP, Matsuzawa Y, Yamashita S, Shai I, et al. Visceral and ectopic fat, atherosclerosis, and cardiometabolic disease: a position statement. Lancet Diabetes Endocrinol. 2019;7:715-25.
19. Soares-Jr JM, Hayashida SAY, Marcondes JAM, Maciel GAR, Barcellos CRG, Maffazioli GDN, et al. Influence of phenotypes on the metabolic syndrome of women with polycystic ovary syndrome over a six-year follow-up in Brazil. Biomedicines. 2023;11:3262.
20. Tao Z, Zuo P, Ma G. Association of weight-adjusted waist index with cardiovascular disease and mortality among metabolic syndrome population. Sci Rep. 2024;14:18684.
21. Martínez-González MA, Gea A, Ruiz-Canela M. The Mediterranean diet and cardiovascular health: a critical review. Circ Res. 2019;124:779-98.


CC BY

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260862

# Fundus camera and artificial intelligence for diabetic retinopathy screening in a São Paulo shelter

Julia Gimenes da Silva Dias[1]*, Gabriel Costa de Andrade[2], João Victor Tenório Lossio de Macedo[2], Andrezza Soldera Graziano[1]

## SUMMARY

**OBJECTIVE:** The aim of this study was to evaluate the applicability of a smartphone-based portable fundus camera for diabetic retinopathy screening in elderly residents of a specialized shelter in São Paulo, Brazil, considering image acquisition by non-ophthalmologist professionals and remote case classification.

**METHODS:** Prospective pilot interventional study at the Centro de Acolhida Especial para Idosos Morada São João. Thirty-eight participants (76 eyes) underwent retinal imaging using the Eyer® portable fundus camera (Phelcom Technologies, Brazil) by trained non-ophthalmologist professionals. Images were remotely evaluated by a retinal specialist and compared to the EyerMaps® artificial intelligence system.

**RESULTS:** Mean age was 71.6±7.0 years; 78.9% male. Fifteen participants (39.5%) had confirmed diabetes mellitus. The specialist identified abnormalities in 70 of 76 eyes (92.1%), most frequently glaucoma suspicion, media opacity suggestive of cataract, chorioretinal changes, and age-related macular degeneration signs; no diabetic retinopathy was identified. Against the specialist's assessment of any ocular abnormality as the reference standard, the artificial intelligence system classified 22 eyes as abnormal, yielding sensitivity 25.7% (95%CI 16.9–37.0), specificity 33.3% (95%CI 9.7–70.0), positive predictive value 81.8% (95%CI 61.5–92.7), negative predictive value 3.7% (95%CI 1.0–12.5), and overall accuracy 26.3% (95%CI 17.7–37.2).

**CONCLUSION:** The Eyer® device demonstrated operational feasibility for ocular screening by trained non-ophthalmologists. The high prevalence of media opacities and non-diabetic retinopathy conditions compromised artificial intelligence performance, highlighting the need for broader screening strategies in elderly and socially vulnerable populations.

**KEYWORDS:** Diabetic retinopathy. Photography. Screening. Artificial intelligence. Telemedicine.

## INTRODUCTION

Diabetes mellitus is one of the most significant public health challenges of our time, with chronic complications directly impacting morbidity, quality of life, and healthcare costs worldwide. Among these, diabetic retinopathy stands out as a leading cause of preventable blindness, affecting over 100 million people globally[1,2].

In Latin America, prevalence ranges from 15.0% in Costa Rica to 32.7% in Brazil—the highest among countries assessed—reflecting disparities in diabetes prevalence and healthcare system inequalities[3]. In Brazil, this is compounded by the uneven distribution of ophthalmological services concentrated in urban referral centers, with limited primary care screening capacity.

Diabetic retinopathy is a progressive retinal complication of chronic hyperglycemia, classified as non-proliferative (mild, moderate, and severe) and proliferative, with diabetic macular edema possible at any stage. Early stages feature microaneurysms, hemorrhages, and hard exudates; advanced ischemia leads to neovascularization, vitreous hemorrhage, and tractional retinal detachment[4,5]. Despite this severity, visual loss is preventable with adequate screening and timely treatment—yet many patients remain asymptomatic until advanced disease[5-7].

Portable fundus cameras and smartphone-based devices enable retinal imaging outside traditional settings, including primary care units and long-term care facilities (LTCFs) for the elderly, and artificial intelligence (AI) systems may reduce dependence on specialist evaluation[6-8]. Barriers are especially

[1]Irmandade da Santa Casa de Misericórdia de São Paulo – São Paulo (SP), Brazil.
[2]Irmandade da Santa Casa de Misericórdia de São Paulo, Department of Ophthalmology – São Paulo (SP), Brazil.
*Corresponding author: juliad912@gmail.com

Conflicts of interest: the Eyer® portable fundus camera and access to the EyerMaps® artificial intelligence system were provided by Phelcom Technologies (São Carlos, Brazil) free of charge for research purposes only. The company had no role in the study design, data collection, analysis, interpretation of results, or the decision to submit the manuscript for publication. No author received any financial compensation or holds any financial interest in Phelcom Technologies. The authors declare no other conflicts of interest. Funding: this study received no specific funding from public, commercial, or non-profit funding agencies. The Eyer® device was made available by Phelcom Technologies for research purposes.
Received on June 12, 2026. Accepted on July 19, 2026.
Scientific Editor: Roseli Nomura

pronounced in elderly and vulnerable populations—functional limitations, comorbidities, and transportation difficulties contribute to underdiagnosis. On-site retinal imaging may overcome these barriers[3,6,7].

This study evaluates the feasibility of the Eyer® portable fundus camera (Phelcom Technologies, São Carlos, Brazil) for diabetic retinopathy screening in elderly residents of the Centro de Acolhida Especial para Idosos Morada São João, with images obtained by trained non-ophthalmologists. Secondarily, it investigates the diagnostic performance of EyerMaps® AI compared to a retinal specialist.

## METHODS

### Study design and population

A prospective pilot interventional study was conducted at the Centro de Acolhida Especial para Idosos Morada São João, central São Paulo, Brazil, between April and September 2025. The study was approved by the Research Ethics Committee of Irmandade da Santa Casa de São Paulo on 30 January 2025 (Opinion No. 7.351.391; CAAE 85488024.7.0000.5479). All study procedures were performed after ethics approval: chart and demographic data were extracted from institutional records from 3 April 2025, and retinal image acquisition was carried out between July and September 2025. Participants were enrolled as a convenience series: all residents aged ≥18 years with a prior diagnosis of diabetes mellitus and/or systemic arterial hypertension—the latter included as a shared cardiovascular risk factor—who were present at the institution during the study period and agreed to sign the Informed Consent Form (ICF) were considered eligible. The aim was to assess imaging feasibility across residents at cardiometabolic risk rather than in diabetic residents exclusively. Individuals under 18 years or who did not consent were excluded. Of 209 residents present during the study period, 38 (18.2%) were enrolled. No a priori sample size calculation was performed given the exploratory nature of the study.

### Procedures and image acquisition

After informed consent, participants were assessed for demographic, clinical, and biochemical data from institutional records. Anterior segment and fundus images of both eyes were acquired using the Eyer® device by non-ophthalmologist professionals (nursing staff, medical students, and physicians without prior ophthalmological experience), all trained by an ophthalmologist. All images were acquired without pharmacological mydriasis (non-mydriatic imaging). Images were uploaded to the EyerCloud® cloud platform (Phelcom Technologies), compliant with the Brazilian General Data Protection Law (*Lei Geral de Proteção de Dados Pessoais* [LGPD]; Law No. 13,709/2018).

### Image evaluation and classification

A retinal specialist remotely classified images as gradable or non-gradable. For gradable images, diabetic retinopathy was categorized as non-referable (absent or mild-to-moderate non-proliferative without maculopathy) or referable (severe non-proliferative, proliferative, or diabetic macular edema). Non-gradable cases had anterior segment images reviewed for media opacities. The specialist performed the evaluation blinded to the EyerMaps® output; conversely, the AI system processed images without access to participants' clinical information. Specialist findings were subsequently compared to EyerMaps® automated analysis.

### Statistical analysis

Categorical variables were presented as absolute and relative frequencies; continuous variables as mean±standard deviation or median (interquartile range). AI diagnostic performance was assessed by sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and overall accuracy, using the specialist's assessment as the reference standard. Ninety-five percent confidence intervals (95%CIs) for each metric were calculated using the Wilson score method.

## RESULTS

### Demographic and clinical characteristics

The flow of participants is illustrated in Figure 1. Thirty-eight participants (76 eyes) were included, predominantly male (78.9%), aged 60 to 90 years (mean 71.6±7.0). Fifteen (39.5%) had confirmed diabetes mellitus, 20 (52.6%) denied it, and 3 (7.9%) were unable to provide information. Table 1 summarizes the main sample characteristics.

### Specialist findings

No adverse events were recorded during image acquisition. Ocular abnormalities were identified in 70 of 76 eyes (92.1%); only 6 eyes (7.9%) from two participants had no abnormal findings. No diabetic retinopathy (referable or non-referable) was identified among gradable eyes; 22 eyes were non-gradable owing to media opacity. Abnormalities were heterogeneous and predominantly unrelated to diabetic retinopathy: glaucoma suspicion or increased cup-to-disc ratio (31.6%), media opacity

suggestive of cataract limiting fundus assessment (28.9%), tessellated and atrophic changes (13.2%), chorioretinal atrophy and scars (10.5%), age-related macular degeneration (AMD) signs (7.9%), disc pallor and peripapillary atrophy (10.5%), and tilted disc or pathological myopia (7.9%). Table 2 details the findings.

## Artificial intelligence system performance

Retinal images were evaluated by the specialist approximately 1 month after acquisition. Using the specialist's assessment of any ocular abnormality as the reference standard, EyerMaps® classified 22 eyes as abnormal and 54 as normal. Among the 70 eyes classified as abnormal by the specialist, the AI correctly

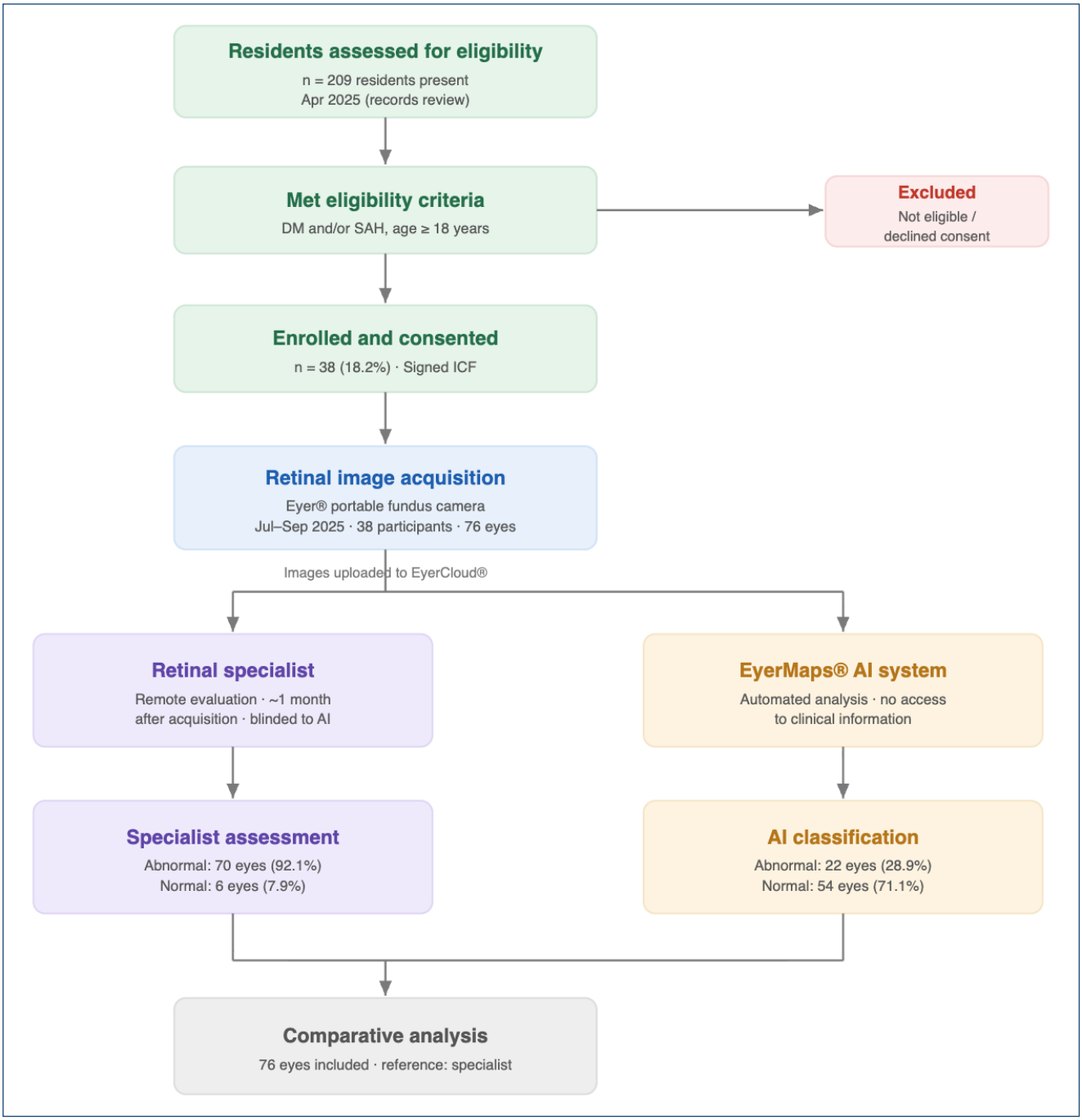



**Figure 1.** Flow of participants, from eligibility assessment to comparative analysis of specialist versus artificial intelligence classification. ICF: Informed Consent Form; DM: diabetes mellitus; SAH: systemic arterial hypertension; AI: artificial intelligence. [First author: Dias JGS—Figure 1].

**Table 1.** Demographic and clinical characteristics of participants (n=38).

| Characteristic | n | % |
|---|---|---|
| Male sex | 30 | 78.9 |
| Age range (years) | 60–90 | – |
| Mean age (years) | 71.6±7.0 | – |
| Diabetes mellitus | | |
| Confirmed | 15 | 39.5 |
| Denied | 20 | 52.6 |
| Unknown | 3 | 7.9 |
| Comorbidities | | |
| Systemic arterial hypertension | 30 | 78.9 |
| Dyslipidemia | 12 | 31.6 |
| Previous cardiovascular disease | 4 | 10.5 |
| Cataract | 12 | 31.6 |
| Glaucoma (prior diagnosis) | 3 | 7.9 |

All 38 medical records were reviewed; comorbidity data are complete, and absence of a listed comorbidity denotes documented absence of that diagnosis. Diabetes status was self-reported and could not be determined in three participants. HbA1c and diabetes duration were unavailable owing to the absence of digitized institutional records.

**Table 2.** Main categories of findings on fundus examination by the specialist (n=76 eyes).

| Category of finding | Eyes (n=76) |
|---|---|
| No abnormalities | 6 (7.9%) |
| Glaucoma suspicion/increased cup-to-disc ratio | 24 (31.6%) |
| Media opacity/cataract (limiting assessment) | 22 (28.9%) |
| Tessellated/senile atrophic changes | 10 (13.2%) |
| Chorioretinal atrophy/scars | 8 (10.5%) |
| AMD/drusen/RPE changes | 6 (7.9%) |
| Disc pallor/peripapillary atrophy | 8 (10.5%) |
| Tilted disc/pathological myopia | 6 (7.9%) |

AMD: age-related macular degeneration; RPE: retinal pigment epithelium. One eye may present multiple findings simultaneously.

flagged 18 (true positives) and missed 52 (false negatives); among the six eyes classified as normal by the specialist, the AI flagged four (false positives) and correctly classified two (true negatives). The system demonstrated sensitivity of 25.7% (95%CI 16.9–37.0), specificity of 33.3% (95%CI 9.7–70.0), PPV of 81.8% (95%CI 61.5–92.7), NPV of 3.7% (95%CI 1.0–12.5), and overall accuracy of 26.3% (95%CI 17.7–37.2). These metrics reflect detection of any ocular abnormality rather than diabetic retinopathy specifically. Despite the very low sensitivity—indicating that the majority of abnormal cases were missed—the PPV of 81.8% suggests that when the system flags an abnormality, the finding is likely genuine, although the wide CI for specificity (based on only six normal eyes) indicates this estimate is imprecise.

## DISCUSSION

This study demonstrated that ocular screening using the Eyer® portable fundus camera by non-ophthalmologist professionals is operationally feasible in an LTCF setting. Images were obtained in all recruited participants without transfer, following brief ophthalmologist-led training, corroborating prior studies validating portable cameras in limited-resource environments[6,7].

The very high prevalence of ocular abnormalities (92.1%) was the most striking finding, with glaucoma suspicion, cataract, atrophic chorioretinal changes, and AMD predominating—conditions unrelated to diabetic retinopathy. This has relevant clinical implications: a substantial proportion of residents may benefit from ophthalmological evaluation regardless of diabetes diagnosis, and LTCF screening strategies should address a broader spectrum of ocular disease. Media opacities compromised image quality in some cases, a challenge inherent to elderly populations and likely accentuated by non-mydriatic acquisition; systematic mydriasis and integration with cataract detection programs may mitigate this in future protocols.

No diabetic retinopathy was identified among gradable eyes. Among the 15 participants with confirmed diabetes, this may be consistent with recent diagnosis and/or adequate metabolic control, in which established retinopathy would not yet be expected. This interpretation remains tentative, as glycated hemoglobin and diabetes duration were unavailable; the small diabetic subgroup, self-reported diabetes status, and media opacities limiting fundus assessment may also account for the absence of detected retinopathy.

EyerMaps® sensitivity (25.7%) was substantially lower than the 80–96% reported in diabetic retinopathy screening validation studies[8]. This reflects a key methodological difference: validation studies were conducted in selected diabetic populations with a focus on referable diabetic retinopathy, whereas we compared AI performance against all ocular findings, including glaucoma, AMD, and chorioretinal atrophies for which the algorithm was not trained. These metrics should therefore be read as the algorithm's performance for detection of any abnormality, not as a validation for its intended indication; rather, they illustrate the risk of applying a disease-specific algorithm as a general screening tool in a population with a high non-diabetic pathology burden. Nevertheless, the PPV of 81.8% preserves its utility as a prioritization tool.

From a public health perspective, these findings suggest that LTCF screening programs should be designed around the broad spectrum of age-related ocular disease rather than diabetic retinopathy alone, paired with structured referral pathways for cataract and glaucoma. Portable, non-mydriatic imaging with remote interpretation is compatible with existing Unified Health System (SUS) primary-care and telehealth workflows, although formal cost and resource evaluation is needed.

Limitations include the absence of digitized records (precluding HbA1c and diabetes duration data), the small sample (38 participants, 76 eyes) with correspondingly wide CIs limiting the precision of diagnostic estimates and generalizability, recruitment failures introducing potential selection bias, self-reported diabetes status, and the methodological limitation of comparing AI performance to all ocular findings rather than diabetic retinopathy alone. Findings may also differ across Brazilian regions and LTCF profiles, and socioeconomic factors likely influence screening uptake. Future studies should include stratified analyses by finding type, structured recruitment workflows, protocols for non-gradable images, and image-quality strategies for media opacity.

## CONCLUSION

The Eyer® portable fundus camera demonstrated operational feasibility for ocular screening by trained non-ophthalmologist professionals at an LTCF, without requiring patient transfer. The high prevalence of ocular abnormalities—predominantly unrelated to diabetic retinopathy—and the resultant compromise of AI performance highlight that LTCF screening strategies must go beyond diabetic retinopathy and be supported by organized referral pathways to specialized care. Portable fundus photography with remote interpretation is a promising tool for broadening ocular diagnostic access in vulnerable populations when implemented within structured, context-aware protocols.

## ACKNOWLEDGMENTS

The authors thank the staff and residents of the Centro de Acolhida Especial para Idosos Morada São João for their participation. The Eyer® device was made available by Phelcom Technologies (São Carlos, Brazil) for research purposes. No external funding was received for this study.

## ETHICS APPROVAL

This study was approved by the Research Ethics Committee (CEP) of the Irmandade da Santa Casa de Misericórdia de São Paulo (Opinion No. 7.351.391; CAAE 85488024.7.0000.5479; approved 30 January 2025), in compliance with CNS Resolution No. 466/2012 and the Declaration of Helsinki.

## AUTHORS' CONTRIBUTIONS

**JGSD:** Conceptualization, Data curation, Formal Analysis, Investigation, Writing – original draft. **GCA:** Conceptualization, Supervision, Validation, Writing – review & editing. **JVTLM:** Data curation, Formal Analysis, Investigation, Writing – original draft. **ASG:** Data curation, Formal Analysis, Investigation, Writing – original draft.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES

1. Tan TE, Wong TY. Diabetic retinopathy: looking forward to 2030. Front Endocrinol (Lausanne). 2023;13:1077669.
2. Teo ZL, Tham YC, Yu M, Chee ML, Rim TH, Cheung N, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. Ophthalmology. 2021;128(11):1580-91.
3. Machado JS, Brandão MN, Castro CT, Ferreira TS, Santos LHPE, Medeiros DS. Prevalence and associated factors of diabetic retinopathy in Latin American countries: a scoping review. Rev Bras Oftalmol. 2024;83:e0034.
4. Gettinger K, Lee D, Tomita Y, Negishi K, Kurihara T. Diabetic retinopathy, a comprehensive overview on pathophysiology and relevant experimental models. Int J Mol Sci. 2025;26(20):9882.
5. Chang TJ. Update in the epidemiology, risk factors, screening, and treatment of diabetic retinopathy. J Diabetes Investig. 2021;12(8):1322-5.
6. Chong DD, Das N, Singh RP. Diabetic retinopathy: screening, prevention, and treatment. Cleve Clin J Med. 2024;91(8):503-11.
7. Vujosevic S, Aldington SJ, Silva P, Hernández C, Scanlon P, Peto T, et al. Screening for diabetic retinopathy: new perspectives and challenges. Lancet Diabetes Endocrinol. 2020;8(4):337-47.
8. Rajesh AE, Davidson OQ, Lee CS, Lee AY. Artificial intelligence and diabetic retinopathy: AI framework, prospective studies, head-to-head validation, and cost-effectiveness. Diabetes Care. 2023;46(10):1728-39.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20261603

# Career, Autonomy, Rights and Equity among Doctors: Brazilian Medical Association survey

Luciana Rodrigues Silva[1,2], Maria Rita de Souza Mesquita[1,3], Claudia Navarro Carvalho Duarte Lemus[1], Alex Jones Flores Cassenote[4,5]*, Mateus de Souza Ribeiro[6], José Eduardo Lutaif Dolci[1,7], Cesar Eduardo Fernandes[1,8]

## SUMMARY

**OBJECTIVE:** This study aimed to describe the design and linkage procedures of the Career, Autonomy, Rights and Equity among Doctors Brazil survey coordinated by the Brazilian Medical Association, and to assess the composition and geographic coverage of its analytic sample against national benchmarks.

**METHODS:** This national cross-sectional online survey was conducted from January 29 to September 10, 2026. The invitation frame came from Demografia Médica no Brasil, a census-based investigation of the physician workforce conducted since 2010. Four nonoverlapping random samples totaling 110,000 physicians were invited; participation was voluntary. Survey records were audited, deduplicated, linked to professional sources through a restricted key, compared with 2024 national benchmarks, and calibrated for sex, region, and age group.

**RESULTS:** Of 4,437 submissions, 4,002 physicians formed the analytic sample. Women accounted for 51.6%, mean age was 48.3 years, and 70.9% had at least one recognized specialty, compared with 49.3%, 44.8 years, and 59.1% in the observed 2024 national benchmark. Calibration weights ranged from 0.396 to 3.134; the Kish weight-based design effect was 1.259, and the effective sample size was 3,180.

**CONCLUSION:** Career, Autonomy, Rights and Equity among Doctors Brazil documents a linked physician sample and quantifies differences from national benchmarks. The low matched participation fraction and voluntary response create substantial potential for selection bias. Random invitations, linkage, and calibration do not establish population representativeness or remove unmeasured nonresponse bias. The findings primarily characterize the analytic sample and inform the interpretation of subsequent thematic studies.

**KEYWORDS:** Physicians. Surveys and questionnaires. Gender equity. Health workforce. Data linkage. Brazil.

## INTRODUCTION

Women are projected to account for more than half of Brazil's physicians in 2026[1]. This numerical majority does not ensure equity in pay, working conditions, specialty choice, leadership, or career progression. Mainardi et al. investigated remuneration; Amaral et al. and Suzuki et al. studied workplace discrimination; Motter et al. examined surgical leadership; and Marhoffer et al. reviewed gender in academic promotion[2-6]. Together, these studies do not provide a shared participant base for comparing several career dimensions under one recruitment and linkage design.

These limitations matter for a voluntary online survey. Random invitation does not make respondents a probability sample; response, linkage, and missingness may change each analytic subset. External census comparisons and calibration clarify the reach and limits of inference but cannot correct unmeasured nonresponse. Reporting these steps supports comparable thematic analyses of gender, work, and income[7-11].

The methodological gap is how to define and evaluate a common analytic universe for focused studies with consistent denominators, linked professional information, and explicit limits to generalization. Career, Autonomy, Rights and Equity among Doctors (CARE-DOC) Brazil links measures of work,

[1]Brazilian Medical Association – São Paulo (SP), Brazil.
[2]Universidade Federal da Bahia – Salvador (BA), Brazil.
[3]Universidade Federal de São Paulo – São Paulo (SP), Brazil.
[4]Universidade de São Paulo, Institute of Advanced Studies – São Paulo (SP), Brazil.
[5]Universidade Brasil – Fernandópolis (SP), Brazil.
[6]Santa Marcelina College, School of Medicine – São Paulo (SP), Brazil.
[7]Santa Casa de São Paulo School of Medical Sciences – São Paulo (SP), Brazil.
[8]ABC Medical School – Santo André (SP), Brazil.
*Corresponding author: cassenote@usp.br
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on October 08, 2026. Accepted on October 14, 2026.
Scientific Editor: Roseli Nomura

family, autonomy, discrimination, leadership, remuneration, and future expectations to professional information. This article describes the instrument, recruitment, processing, and linked sample, and empirically assesses deviations from national benchmarks, geographic coverage, and calibration-weight dispersion. Its contribution is descriptive and methodological; the potential for future thematic research is a secondary application.

## METHODS

### Study design and instrument

CARE-DOC Brazil was a national cross-sectional online survey coordinated by the Brazilian Medical Association from January 29 to September 10, 2026. The Portuguese questionnaire was developed through literature review, institutional discussion, and version review, drawing on Brazilian physician studies[1,2]. It combined single- and multiple-choice, numeric, and five-point agreement items across work, family, labor rights, discrimination, leadership, gender attitudes, and future plans. For up to seven regular medical work relationships, the repeatable module recorded care setting, financing (public Brazilian Unified Health System, insurance, or direct payment), contract, leadership, activity, weekly hours, and monthly net-income band. Contracts included employment under Brazilian labor law (Consolidation of Labor Laws), statutory public service, contracting through a legal entity (pessoa jurídica, PJ), independent practice, and residency or fellowship stipends. Twenty selected Ambivalent Sexism Inventory statements and four modified work–family conflict statements used five-point agreement responses[12,13]. Their wording and response formats were not assumed equivalent to the source instruments. Formal cognitive testing and prefield psychometric validation were not conducted; any composite score requires separate evaluation.

Reporting followed Consensus-Based Checklist for Reporting of Survey Studies, Checklist for Reporting Results of Internet E-Surveys, and Strengthening the Reporting of Observational Studies in Epidemiology guidance[7,8,14]. Figure 1 presents the audited sample flow.

### Recruitment, linkage, and privacy

The invitation frame came from Demografia Médica no Brasil, a nationwide census-based investigation conducted since 2010[1]. Four nonoverlapping random invitation waves of 10,000, 20,000, 30,000, and 50,000 physicians totaled 110,000; participation was voluntary. The Brazilian Medical Association coordinated mobilization, not the frame. Among 4,437 submissions, we kept the more complete of repeated federative unit-Regional Council of Medicine registration pair records (most recent if tied), removing 230. Seven initially had no male/female sex classification. After a provisional geographic balancing step, 148 women were randomly excluded proportionally across adjusted-state strata from the remaining 4,200 (seed 20260919) to reduce the female excess, leaving 4,052. We then excluded 33 without a recoverable linkage key and nine repeated individual linkages. Six of the seven sex-unclassified records were recovered from administrative data; one remained excluded. Finally, 14 women (12 in São Paulo, one each in Minas Gerais and Rio Grande do Sul) were randomly excluded within provisional state strata (seed 20260921), yielding 4,002. Of these, 3,939 matched the invitation frame and 63 did not. Before analysis, we compared excluded and retained self-reported women on raw questionnaire characteristics.

Deduplication and deterministic linkage used a restricted key available only to the data-management team. Linked sources supplied or verified age, graduation, school sector,

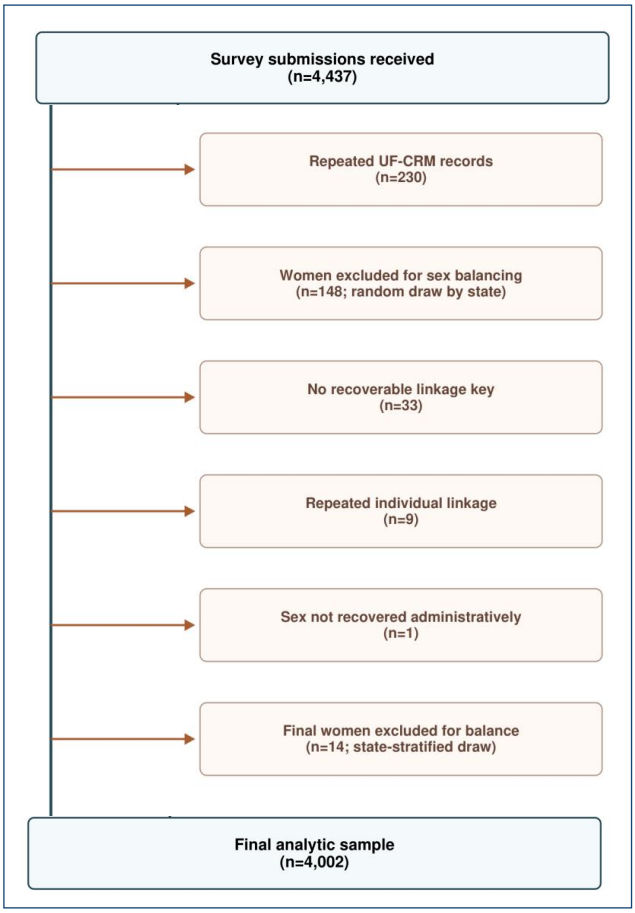



**Figure 1.** Formation of the final analytic sample. The flow records all exclusions from 4,437 submissions to 4,002 participants. Direct identifiers were excluded from analytic outputs.

specialty and area-of-practice records, residency, and specialty-board certification. Registry data corrected the analysis sex classification for 45 records; original self-reports were retained separately. Direct identifiers were excluded from analysis. Absence from a title registry was coded as no recorded specialty or practice area.

### Study size and participation

The analytic sample included consenting physicians with a unique recoverable linkage key and sex; matching the invitation list was not required. As a descriptive precision benchmark, the finite-population formula with 597,428 physicians in 2024, 50% expected prevalence, 95% confidence, and a two-percentage-point margin under simple random sampling yielded 2,392 physicians[15]. This benchmark does not establish achieved precision, population representativeness, or power for thematic outcomes. Of 110,000 selected invitees, 3,939 appeared in the analytic sample (3.6%). This matched participation fraction is not an American Association for Public Opinion Research response rate: invitation delivery and eligibility among nonrespondents were unverified[16]. The final analytic sample also included 63 physicians without a match to the invitation frame. Thus, dividing all 4,002 participants by 110,000 invitees (3.64%) does not estimate participation among sampled invitees. Neither fraction establishes that respondents resemble nonrespondents.

### Population comparisons and statistical analysis

The principal benchmark was Demografia Médica no Brasil 2025, which describes physicians and active registrations in 2024[1]. Sex, mean age, and specialty were compared with individual-physician estimates. Regional comparisons used participants' original reported state and active registrations, which may include more than one registration per physician. The provisional balancing file reassigned 103 retained participants to other state strata; final geography and weights use their original state. School sector and time since graduation came from the linked administrative extract dated October 10, 2024. Age and time since graduation were standardized to December 31, 2024. For the municipality map, a restricted key linked respondents to the 2024 physician registry. We selected an exact state-and-registration match or an unambiguous municipality within the original state, plotted aggregate counts within Brazilian Institute of Geography and Statistics (IBGE) municipal boundaries, and omitted ambiguous addresses. Registry correspondence address may differ from residence or practice location.

We calculated Wilson 95%CIs for proportions, a t-based interval for mean age, standardized differences, and state observed-to-expected ratios (original-state survey share divided by the national active-registration share). Pearson goodness-of-fit tests assessed sex, five regions, and age groups (<35, 35–44, 45–54, 55–64, and ≥65 years). Intervals and p-values assume simple random sampling and do not account for voluntary response or post-collection selection. Calibration weights were estimated by iterative proportional fitting to 2024 sex totals, age groups among physicians with known age, and regional active-registration totals, using the original state[9]. Seven missing ages were assigned the sex-specific sample median only for calibration; descriptive age used 3,995 observed values. Weights were normalized to mean one. Descriptive estimates used available cases and variable-specific denominators. We reported weight range, coefficient of variation (CV), Kish weight-based design effect (DEFF)=1+CV$^2$, and effective n=n/DEFF[17]. This DEFF describes weight dispersion, not the full sampling variance or nonresponse bias. The effective sample size is therefore a weight-based information benchmark, not evidence of unbiased population estimates. Calibration aligns the selected observed margins but cannot establish exchangeability between respondents and nonrespondents or correct selection related to unmeasured characteristics. Analyses used R version 4.6.1 (R Foundation for Statistical Computing, Vienna, Austria).

### Ethics

The Research Ethics Committee of the University of São Paulo School of Medicine approved the study (Certificate of Presentation for Ethical Consideration 71626323.8.0000.0068). Participation was voluntary and electronic informed consent was obtained before questionnaire completion.

## RESULTS

### Analytic sample and external comparison

The final sample comprised 4,002 physicians (90.2% of submissions), including 3,939 matched to the invitation frame (3.58% of 110,000 invitees) and 63 without a frame match. Mean questionnaire completeness was 99.0% among included participants; it does not indicate high participation among invitees. Women represented 51.6% (95%CI 50.1–53.2). Mean age was 48.3 years (standard deviation [SD] 14.3; 95%CI 47.9–48.7), 3.5 years above the national mean. At least one recognized specialty was recorded for 70.9% of participants versus 59.1% nationally. The sample also contained more public-school graduates and longer-term graduates (Table 1).

**Table 1.** Participant profile and comparison with national physician benchmarks.

| Characteristic | Survey n/N | Estimate | 95%CI | DMB 2024 | Difference |
|---|---|---|---|---|---|
| Women | 2,066/4,002 | 51.6% | 50.1–53.2 | 49.3% | +2.3 pp |
| Age, mean (SD), years | 3,995/4,002 | 48.3 (14.3) | 47.9–48.7 | 44.8 (15.7) | +3.5 years |
| Black or brown | 653/3,985 | 16.4% | 15.3–17.6 | – | – |
| North | 130/4,002 | 3.2% | 2.7–3.8 | 4.8% | -1.6 pp |
| Northeast | 568/4,002 | 14.2% | 13.1–15.3 | 19.3% | -5.1 pp |
| Southeast | 2,209/4,002 | 55.2% | 53.7–56.7 | 51.1% | +4.1 pp |
| South | 788/4,002 | 19.7% | 18.5–21.0 | 15.8% | +3.9 pp |
| Central-West | 307/4,002 | 7.7% | 6.9–8.5 | 9.0% | -1.3 pp |
| Public medical school | 2,244/3,887 | 57.7% | 56.2–59.3 | 53.7% | +4.1 pp |
| No recognized specialty | 1,163/4,002 | 29.1% | 27.7–30.5 | 40.9% | -11.8 pp |
| One recognized specialty | 1,858/4,002 | 46.4% | 44.9–48.0 | 40.9% | +5.5 pp |
| Two recognized specialties | 821/4,002 | 20.5% | 19.3–21.8 | 16.0% | +4.5 pp |
| Three or more specialties | 160/4,002 | 4.0% | 3.4–4.7 | 2.2% | +1.8 pp |

CI: confidence interval; DMB: Demografia Médica no Brasil; pp: percentage points; SD: standard deviation. Regional benchmarks use active registrations. A dash indicates no comparator.

### Geographic coverage and calibration

All 27 federative units were represented. The North and Northeast were underrepresented by 1.6 and 5.1 percentage points, whereas the Southeast and South were overrepresented by 4.1 and 3.9 points. Goodness-of-fit tests identified differences for sex (chi-square[1]=8.65; p=0.003), original-state region (chi-square[4]=126.84; p<0.001), and age group (chi-square[4]=799.04; p<0.001). Raking converged in 25 iterations. Weights had mean 1.000, CV 0.508, range 0.396–3.134, and 1st–99th percentiles 0.396–2.493. The Kish weight-based DEFF was 1.259 and effective sample size was 3,180. Among 162 excluded versus 2,073 retained self-reported women, 38.3% versus 36.2% reported at least three jobs, and 20.4% versus 17.3% reported Black/Brown race or color.

For 3,928 of 4,002 participants (98.2%), the correspondence-address municipality could be linked unambiguously within the original survey state; they represented 648 municipalities. The remaining 74 were not mapped (50 registry records only in another state, 20 without a municipality code, and four without a registry match). Panel B compares each original state's share of the survey with its share of active physician registrations in the national benchmark (Figure 2).

## DISCUSSION

CARE-DOC Brazil documents the construction and composition of a linked physician sample and demonstrates measurable deviations from national benchmarks. Respondents were older and more frequently specialized, and their regional distribution differed from the benchmark. The calibration diagnostics quantify the dispersion of weights used to align selected margins. These are descriptive methodological findings about the observed sample. The linked database may support thematic studies, but that future utility is not itself evidence of substantive findings on working conditions, gender equity, or health-service outcomes in the present article.

The principal limitation is the very low matched participation fraction: only 3,939 of 110,000 invited physicians entered the analytic sample. Random selection strengthens the invitation framework but does not ensure a representative final sample when participation is voluntary. Respondents may differ systematically from nonrespondents in professional engagement, available time, experiences of discrimination, work–family conflict, or interest in the survey topic. These are plausible selection mechanisms, not differences demonstrated by our data. Their direction and magnitude cannot be determined from the available respondent data, and substantial nonresponse selection bias remains possible[10]. Mobilization through specialty groups may also have contributed to the older and more specialized respondent profile, but this cannot be quantified.

Calibration aligned sex, age-group, and regional margins with the chosen 2024 benchmarks[9]. This improves comparability on those observed dimensions but cannot eliminate bias associated with unmeasured determinants of participation or outcomes. Regional targets based on active registrations also differ from individual-physician totals, and the benchmarks

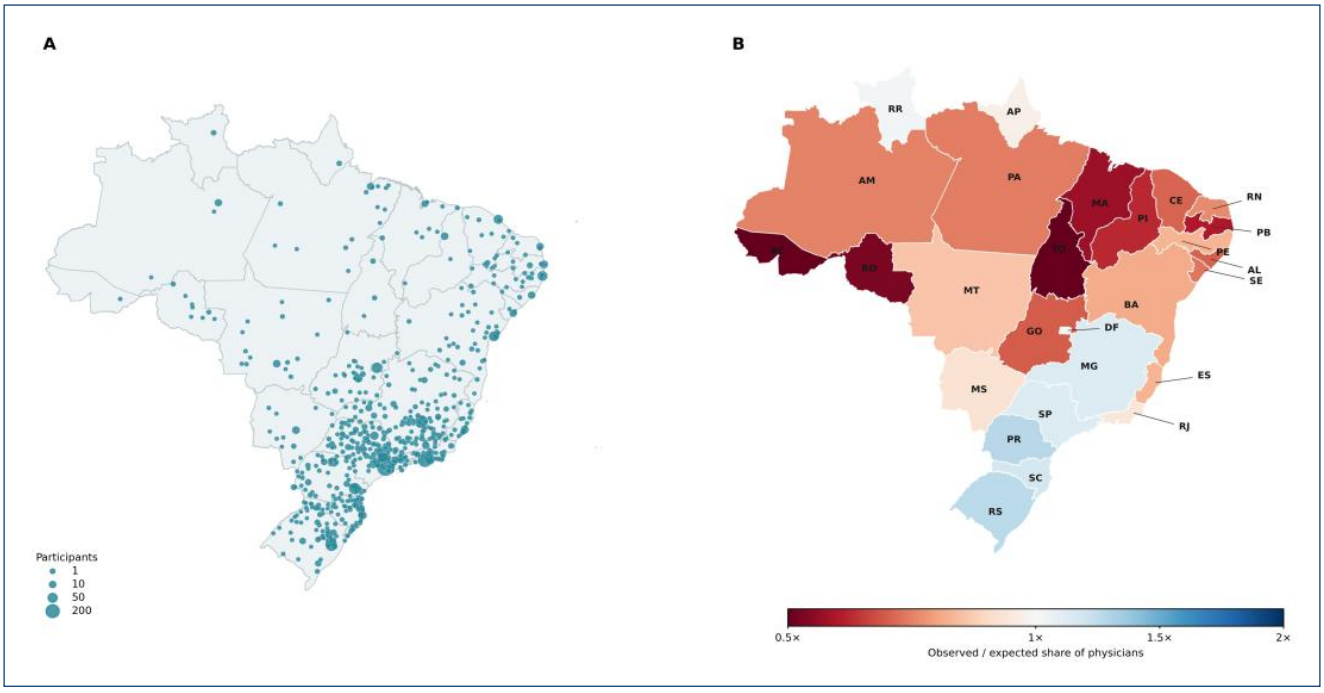



**Figure 2.** Geographic distribution of linked participants and state-level relative participation. (A) Aggregate number of linked participants by municipality of correspondence address in the October 10, 2024 physician registry, plotted within 2024 Brazilian Institute of Geography and Statistics municipal boundaries; 3,928 participants in 648 municipalities. The address does not necessarily indicate residence or practice location. (B) Observed-to-expected ratios using participants' original state; values above 1 indicate a larger survey share than the national share. The color scale is capped at 0.5 and 2.0. State boundaries: IBGE 2024. The Demografia Médica no Brasil 2025 benchmark reports active registrations in 2024.

precede the survey by two years. The Kish weight-based DEFF of 1.259 describes weight dispersion and does not measure total uncertainty or residual selection bias. Thematic studies should present unweighted estimates alongside calibrated sensitivity analyses and report weight dispersion and effective sample size. Agreement between weighted and unweighted estimates would show robustness to the specified calibration, not proof of population validity. Linkage reduced respondent burden and supported classification of career attributes, but it can also generate selection and classification error[11]. Analyses should retain match indicators, document registry rules, and compare linked and unlinked records when relevant. Adapted multi-item constructs require reliability and dimensional assessment before composite scores are compared; the survey as a whole is not a diagnostic or psychometric scale.

Additional limitations include post-collection exclusion of 162 women to balance sex and state, 63 participants unmatched to the invitation frame, differences between 2024 registry data and 2026 responses, self-report error, and no formal prefield cognitive interviewing. Post-collection exclusions may further alter the composition of the sample and thematic estimates. High questionnaire completeness and audited linkage improve data quality among included participants but do not resolve selection into that sample. Representation of all states establishes geographic presence, not representative coverage of physicians within each state. The predominantly descriptive design limits the scientific contribution to documenting sample construction and assessing its observed composition and inferential constraints. It does not estimate causal effects or demonstrate future consequences for physicians or health services. Population prevalence estimates and associations derived from this resource require explicit qualification because calibration cannot justify unrestricted national inference.

## CONCLUSION

CARE-DOC Brazil provides a documented linked analytic sample of 4,002 physicians and an empirical assessment of its composition, geographic presence, and calibration diagnostics. The low matched participation fraction, voluntary response, and post-collection selection limit generalizability. Random invitation and administrative linkage do not overcome these limitations, and calibration cannot remove unmeasured nonresponse bias. Subsequent thematic studies can use this resource to examine patterns among participants, with calibrated sensitivity analyses and explicit limits on population interpretation.

## AUTHORS' CONTRIBUTIONS

**LRS:** Conceptualization, Supervision, Writing – review & editing. **MRSM:** Conceptualization, Writing – review & editing. **CNCDL:** Conceptualization, Writing – review & editing. **AJFC:** Conceptualization, Data curation, Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. **MSR:** Formal Analysis, Methodology, Visualization, Writing – original draft, Writing – review & editing. **JELD:** Conceptualization, Writing – review & editing. **CEF:** Conceptualization, Writing – review & editing.

## DATA AVAILABILITY STATEMENT

De-identified data may be made available from the corresponding author upon reasonable request, subject to applicable ethical and confidentiality restrictions and institutional authorization. Identifiable data will not be shared.

## REFERENCES


1. Scheffer M, coordinator. Demografia Médica no Brasil 2025. Brasília (DF): Ministério da Saúde; 2025.
2. Mainardi GM, Cassenote AJF, Guilloux AGA, Miotto BA, Scheffer MC. What explains wage differences between male and female Brazilian physicians? A cross-sectional nationwide study. BMJ Open. 2019;9(4):e023811. https://doi.org/10.1136/bmjopen-2018-023811
3. Amaral A, Calcado I, Gomez A, Ricci C, Oberlohr V, Mackechnie MC, et al. The perspective of Brazilian women orthopaedic surgeons on gender discrimination: initial insights to understand gender bias in the Brazilian Healthcare System. Cureus. 2024;16(5):e61325. https://doi.org/10.7759/cureus.61325
4. Suzuki DA, Gaui MFD, Rosa DD, Machado KK, Moura FC, Gelatti ACZ, et al. Gender equity and workplace mistreatment in oncology: results from a survey by the Brazilian Society of Clinical Oncology. JCO Glob Oncol. 2025;11:e2400323. https://doi.org/10.1200/GO-24-00323
5. Motter SB, Brandão GR, Iaroseski J, Spadoa JL, Alves AV, Assis Brasil CM, et al. Women representation in academic and leadership positions in surgery in Brazil. Am J Surg. 2022;223(1):71-5. https://doi.org/10.1016/j.amjsurg.2021.07.023
6. Marhoffer EA, Ein-Alshaeba S, Grimshaw AA, Holleck JL, Rudikoff B, Bastian LA, et al. Gender disparity in full professor rank among academic physicians: a systematic review and meta-analysis. Acad Med. 2024;99(7):801-9. https://doi.org/10.1097/ACM.0000000000005695
7. Sharma A, Minh Duc NT, Luu Lam Thang T, Nam NH, Ng SJ, Abbas KS, et al. A consensus-based checklist for reporting of survey studies (CROSS). J Gen Intern Med. 2021;36(10):3179-87. https://doi.org/10.1007/s11606-021-06737-1
8. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res. 2004;6:e34. https://doi.org/10.2196/jmir.6.3.e34
9. Deville JC, Sarndal CE. Calibration estimators in survey sampling. J Am Stat Assoc. 1992;87:376-82. https://doi.org/10.1080/01621459.1992.10475217
10. Rhodes CA, Hu X, Freeman RB, Agrawal R, Cherot E, Dardarian TS, et al. Who's not talking? Nonresponse bias in healthcare employee well-being surveys. J Healthc Manag. 2025;70(5):337-53. https://doi.org/10.1097/JHM-D-24-00166
11. Elstad M, Ahmed S, Røislien J, Douiri A. Evaluation of the reported data linkage process and associated quality issues for linked routinely collected healthcare data in multimorbidity research: a systematic methodology review. BMJ Open. 2023;13(5):e069212. https://doi.org/10.1136/bmjopen-2022-069212
12. Glick P, Fiske ST. The Ambivalent Sexism Inventory. J Pers Soc Psychol. 1996;70:491-512. https://doi.org/10.1037/0022-3514.70.3.491
13. Breyer B, Bluemke M. Work-Family Conflict Scale (ISSP). ZIS - GESIS Leibniz Institute for the Social Sciences; 2016. https://doi.org/10.6102/zis243
14. Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. PLoS Med. 2007;4(10):e296. https://doi.org/10.1371/journal.pmed.0040296
15. Ahmed SK. How to choose a sampling technique and determine sample size for research: a simplified guide for researchers. Oral Oncol Rep. 2024;12:100662. https://doi.org/10.1016/j.oor.2024.100662
16. American Association for Public Opinion Research. Standard definitions: final dispositions of case codes and outcome rates for surveys. 10th ed. Alexandria (VA): AAPOR; 2023.
17. Kish L. Survey sampling. New York (NY): Wiley; 1965.

ORIGINAL ARTICLE

https://doi.org/10.1590/1806-9282.20260403

# Pain, anxiety, and comfort among women undergoing medical termination

Seda Karaçay Yikar[1]*, Melike Nur Sönmez[2], Fikriye Işıl Adigüzel[3], Evşen Nazik[1]

## SUMMARY

**OBJECTIVE:** This study aimed to determine pre- and post-procedural anxiety, intra- and post-procedural pain, and post-procedural comfort among women undergoing medical termination.

**METHODS:** This descriptive correlational study included 55 women who underwent medical termination. Data were collected between April and December 2024 using a Personal Information Form, the Visual Analog Scale, the State-Trait Anxiety Inventory, and the Perianesthesia Comfort Scale.

**RESULTS:** The mean Visual Analog Scale score was 8.96±1.08 during the procedure and 4.81±1.59 after the procedure, with the difference being statistically significant ($p<0.05$). The mean pre-procedural Trait Anxiety Inventory score was 47.34±4.99. The mean pre-procedural State Anxiety Inventory score was 37.85±3.97, whereas the mean post-procedural State Anxiety Inventory score was 36.18±3.42, indicating mild anxiety. The difference between the pre- and post-procedural State Anxiety Inventory scores was statistically significant. The mean post-procedural Perianesthesia Comfort Scale score was 92.52±10.41. Multiple linear regression analysis showed that pain, anxiety, indication for termination, and the need for psychological support collectively explained 34% of the variance in comfort.

**CONCLUSION:** Women undergoing medical termination experienced anxiety both before and after the procedure, severe intra-procedural pain, and moderate post-procedural pain. Pain, anxiety, termination indication, and need for psychological support significantly predicted post-procedural comfort. Medical termination care should adopt a holistic approach integrating effective pain and anxiety management to improve women's comfort.

**KEYWORDS:** Abortion, legal. Anxiety. Pain. Comfort. Nurses.

## INTRODUCTION

Pregnancy is generally expected to progress normally and culminate in childbirth. However, certain complications arising during pregnancy may impair the well-being of both the mother and the fetus[1]. Medical termination may be required when maternal life is threatened or serious fetal anomalies are identified[2,3].

Advances in prenatal screening have increased the detection of fetal anomalies. The incidence of fetal anomalies leading to intrauterine pregnancy loss and medical termination has reached up to 4–5%[4]. The diagnosis of fetal anomalies often obliges families to make the difficult decision of whether to terminate the pregnancy[5]. Medical termination is often experienced as a life crisis associated with psychological, social, and spiritual challenges[2,6]. Recent studies have shown that the medical termination process causes considerable psychological distress in women and significantly increases anxiety levels[7,8].

Holistic care addressing women's comfort improves healthcare outcomes and patient satisfaction. Emotional fluctuations experienced during this process, along with fear, anxiety, and pain experienced during and after the procedure, directly affect women's comfort levels[9]. Therefore, this study aimed to evaluate pre- and post-procedural anxiety, intra- and post-procedural pain, and comfort levels among women undergoing medical termination.

## METHODS

### Study design and participants

This cross-sectional, descriptive study was conducted between April and December 2024 at a city hospital in a province in the Mediterranean region of Turkey. The study included women aged 18–49 years who were admitted for medical termination, had a medical indication, had a gestational age of less than 18 weeks, and had no communication difficulties, psychiatric disorders, or psychiatric medication use.

For the sample size calculation, effect sizes reported in previous studies evaluating pain levels associated with medical termination were used as a reference[10]. A moderate-to-large effect

[1]Çukurova University, Faculty of Health Sciences, Department of Obstetrics and Gynecology Nursing – Adana, Türkiye.
[2]Tokat Gaziosmanpaşa University, Erbaa Faculty of Health Sciences, Department of Women's Health and Diseases Nursing – Tokat, Türkiye.
[3]Medical Park Hospital, Department of Obstetrics and Gynecology – Adana, Türkiye.
*Corresponding author: sedakrcyyikar@gmail.com
Conflicts of interest: the authors declare there is no conflicts of interest. Funding: none.
Received on June 07, 2026. Accepted on July 07, 2026.
Scientific Editor: José Maria Soares Júnior

size (d=0.60) was assumed. With the Type I error probability set at α=0.05 and a sample size of 55, the statistical power to detect changes in pain levels was calculated as 1 - β=0.94.

## Data collection tools

The Personal Information Form included eight researcher-developed questions[7-10].

The Visual Analog Scale (VAS) consists of a 10-cm line displayed horizontally or vertically. One end of the line represents "no pain," whereas the other represents "the worst imaginable pain." Patients are asked to mark the point on the line that best reflects the intensity of the pain they experienced. The distance between the "no pain" anchor and the patient's mark represents the individual's pain intensity[11].

The State-Trait Anxiety Inventory (STAI) is a self-report instrument consisting of 40 items and two subscales: the State Anxiety Inventory (SAI; items 1–20), which assesses situational anxiety, and the Trait Anxiety Inventory (TAI; items 21–40), which measures long-term anxiety tendencies. Scores for each subscale range from 20 to 80, with higher scores indicating greater levels of anxiety; scores above 42 indicate high anxiety[12]. The Cronbach's alpha coefficient in this study was 0.88.

The Perianesthesia Comfort Scale (PCS) consists of 24 items that evaluate individuals' cognitive processes and self-perceptions following invasive procedures. The scale is a six-point Likert-type instrument (1=strongly disagree, 6=strongly agree). Total scores range from 24 to 144, with higher scores indicating greater comfort. The validity and reliability of the scale were established by Üstündağ and Aslan, who reported a Cronbach's alpha coefficient of 0.89[13]. The Cronbach's alpha coefficient for the PCS in this study was 0.80.

## Data collection

The study was conducted between April and December 2024 among women hospitalized in the obstetrics and perinatology service for medical termination. Before the procedure, participants completed the Personal Information Form and STAI, including the SAI and TAI subscales. VAS was administered during the procedure and again 15–20 min afterward. Before discharge, participants completed the SAI and the PCS. A standard medication protocol for cervical effacement and dilation standard (Figure 1) and analgesia protocol for pain management (Figure 2) were implemented for all participants.

| Week of Pregnancy | Goal / Clinical Objective | Misoprostol Dosage | Route of Administration | Misoprostol Administration Interval |
|---|---|---|---|---|
| **Group 1: ≤12 weeks (Early Pregnancy)** | To relax the cervix and facilitate dilation prior to a surgical procedure (vacuum aspiration). | 200 µg (Single Dose) | BU/ SL/ PV* | Misoprostol 800µg every 3 hours until expulsion |
| **Group 2: 13-17 weeks (End of the first trimester / Beginning of the second trimester)** | To ensure deep ripening in order to reduce the risk of cervical trauma and perforation prior to D&E (Dilation and Evacuation). | 200 µg (Single Dose) | BU/ SL/ PV* Preferably; PV/ SL | Misoprostol 400 µg every 3; 200 µg PV* 200 µg SL hours until expulsion |

***Buccal (BU) | Sublingual (SL) |Per Vagina (PV)**

https://www.figo.org/sites/default/files/2024-01/FIGO-Miso-Chart-11-2023%20FINAL%20%281%29%20%281%29.pdf

**Notes on Table and Clinical Monitoring Parameters**

- Clinical Monitoring: Following administration of the medication, patients were closely monitored for potential side effects of misoprostol, such as cramps, vaginal bleeding and shivering/fever, until the time of the procedure.
- Surgical Stage: Once cervical effacement (100 per cent) and dilation (10 cm) were complete in women who had received misoprostol, they were taken to the examination room for the curettage procedure, where the surgical procedure was performed.

**Figure 1.** Misoprostol protocol by gestational week.

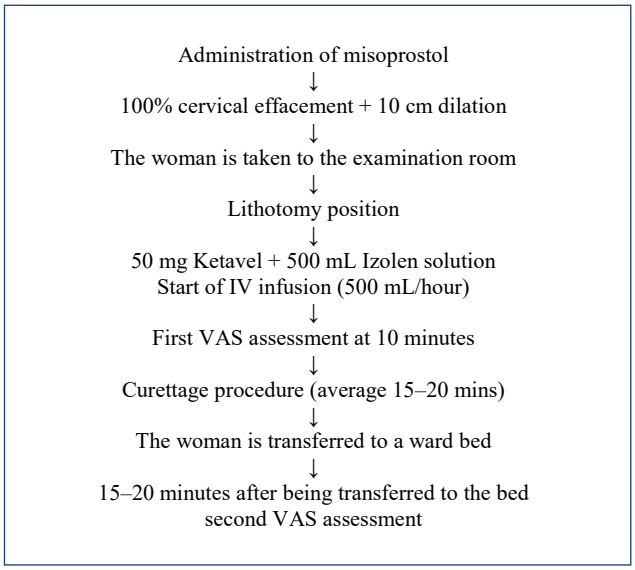



**Figure 2.** Clinical procedure, administration of analgesia and pain assesement procedure.

## Data analysis

Data were analyzed using IBM SPSS Statistics Version 21.0. Normality was assessed using skewness and kurtosis values. Descriptive statistics, dependent-samples t-test, and multiple linear regression analysis were performed. Statistical significance was set at $p<0.05$.

## Ethical considerations

Ethical approval was obtained from the relevant ethics committee (Approval number: 135-41; July 14, 2023). Each patient's written consent was obtained in accordance with the Declaration of Helsinki.

# RESULTS

The mean age of the women was 29.60±6.68 years, and the mean gestational age was 12.78±2.15 weeks (Table 1).

The mean VAS score was 8.96±1.08 during the procedure and 4.81±1.59 15 min afterward. This difference was statistically significant (p<0.05). The mean pre-procedural TAI score was 47.34±4.99, indicating high anxiety levels. The mean SAI score was 37.85±3.97 before the procedure and 36.18±3.42 after the procedure, indicating mild anxiety at both time points. The difference between the pre- and post-procedural SAI scores was statistically significant (p<0.05). The mean post-procedural PCS score was 92.52±10.41, indicating a moderate level of comfort (Table 2).

According to the multiple linear regression analysis, the model predicting PCS scores was statistically significant (R=0.590, $R^2$=0.348, Adjusted $R^2$=0.251, F=3.581, p=0.004), and the

**Table 1.** Distribution of the participants' descriptive characteristics (n=55).

| Characteristics | n | % |
|---|---|---|
| Educational status | | |
| Primary | 26 | 47.3 |
| Secondary | 13 | 23.6 |
| Associate degree or higher | 16 | 29.1 |
| History of curettage/dilation | | |
| Yes | 20 | 36.4 |
| No | 35 | 63.6 |
| Emotion experienced after the procedure | | |
| Sadness or grief | 28 | 50.9 |
| Guilt, regret, or feelings of sinfulness | 10 | 18.2 |
| Anxiety, stress, or fear | 9 | 16.4 |
| Anger or resentment | 5 | 9.1 |
| Relief | 3 | 5.5 |
| Need for psychological support | | |
| Yes | 14 | 25.5 |
| No | 41 | 74.5 |
| Indication for termination | | |
| Intrauterine death | 27 | 49.1 |
| Other indications (hereditary disorders, developmental abnormalities, etc.) | 28 | 50.9 |
| Mean±SD | | |
| Age: 29.60±6.68 (min.:18, max.:43) | | |
| Total number of pregnancies: 3.07±1.77 (min.:1, max.:9) | | |
| Gestational age (weeks): 12.78±2.15 (min.: 8, max.:17) | | |

SD: standard deviation.

**Table 2.** Distribution of women's mean Visual Analog Scale, State-Trait Anxiety Inventory, and Perianesthesia Comfort Scale scores across procedural phases.

| Time | Mean±SD | Min–max | t | p |
|---|---|---|---|---|
| Intra-procedural VAS | 8.96±1.08 | 7–10 | 26.47 | **0.001** |
| Post-procedural VAS | 4.81±1.59 | 1–8 | | |
| Pre-procedural TAI | 47.34±4.99 | 39–57 | | |
| Pre-procedural SAI | 37.85±3.97 | 29–46 | 3.15 | **0.003** |
| Post-procedural SAI | 36.18±3.42 | 30–45 | | |
| Post-procedural PCS | 92.52±10.41 | 24–144 | | |

SD: standard deviation; VAS: Visual Analog Scale; TAI: Trait Anxiety Inventory; SAI: State Anxiety Inventory; PCS: Perianesthesia Comfort Scale. Bold values indicate statistically significant results (p<0.05).

independent variables explained 34% of the variance in comfort scores. The analysis revealed that post-procedural pain at 15 min (β=0.489), indication for termination (β=0.251), and the need for psychological support (β=0.291) were significant predictors of comfort (Table 3).

## DISCUSSION

When pregnancy loss occurs due to fetal health problems, it may lead to psychological difficulties for women. Studies have shown that women experience feelings of anger, resentment, guilt, anxiety, hopelessness, and depression following medical termination[3,14,15]. The findings demonstrated that most women did not seek psychological help. This may be because they were unaware of their need for psychological support.

Women experienced severe pain during the procedure, followed by a significant reduction in pain intensity after the procedure. These findings are consistent with previous studies[16]. The high level of pain experienced during the procedure may be attributable to the invasive nature of the intervention, including cervical dilation and uterine manipulation. The reduction in pain following the completion of the termination procedure is physiologically plausible because uterine stimulation decreases once the intervention has ended. However, post-procedural pain may have been influenced by the timing of the assessment, which occurred before the analgesic reached its maximum effect. In addition, psychological stress may have heightened women's pain perception.

The sudden decision to undergo medical termination, uncertainty about the procedure, fear of pain, inadequate information, concerns about complications, and fetal loss may all contribute to women's anxiety[14,17]. Pre-procedural anxiety scores were significantly higher than post-procedural scores, consistent with previous studies[18-20]. Higher pre-procedural anxiety may reflect uncertainty and fear, whereas the decline after the procedure likely reflects relief following its completion.

According to the model established in the study, pain, anxiety, indications for termination, and the need for psychological support were significant factors affecting women's post-procedural comfort. Women's experience of comfort is a complex and multidimensional phenomenon. The finding that comfort levels decreased as anxiety increased suggests a strong and clinically meaningful association between these variables. Although few studies have directly examined comfort following medical termination, findings from studies involving other patient populations have similarly shown that pain and anxiety negatively affect comfort, supporting the present results[21,22].

These findings indicate that anxiety is an important clinical determinant of comfort during medical termination, highlighting the need for care that addresses both physical and psychological needs. The association between higher anxiety and lower comfort suggests that interventions should extend beyond physical symptom management. Although women reported moderate post-procedural comfort, this does not necessarily reflect optimal well-being, as comfort is a multidimensional concept encompassing physical, psychological, environmental, and sociocultural dimensions. Reduced pain and the resolution of uncertainty may have improved physical comfort, whereas the emotional burden of pregnancy loss likely limited overall comfort.

### Limitations of the study

Conducting the study at a single center was one of the main limitations of this study. The administration of the scales before discharge and the absence of repeated measurements may have limited the evaluation of time-dependent changes in anxiety,

**Table 3.** Multiple linear regression analysis predicting patient comfort.

| Variable | B | Std. error | Beta | t | p | Tolerance | VIF |
|---|---|---|---|---|---|---|---|
| Constant | 105.393 | 19.645 | - | 5.365 | 0.000 | - | - |
| Intra-procedural VAS | -3.171 | 1.602 | -0.331 | -1.979 | 0.054 | 0.496 | 2.017 |
| Post-procedural VAS (15 min) | 3.185 | 1.083 | 0.489 | 2.941 | **0.005** | 0.502 | 1.993 |
| Post-procedural SAI | -0.548 | 0.385 | -0.180 | -1.421 | 0.162 | 0.863 | 1.159 |
| Indication for termination | 1.297 | 0.637 | 0.251 | 2.037 | **0.047** | 0.913 | 1.096 |
| Gestational age (weeks) | -0.015 | 0.593 | -0.003 | -.026 | 0.980 | 0.920 | 1.087 |
| Need for psychological support | 6.903 | 3.134 | 0.291 | 2.203 | **0.033** | 0.828 | 1.208 |
| History of curettage/ dilation | 1.300 | 2.612 | 0.061 | 0.498 | 0.621 | 0.937 | 1.068 |

$R=0.590$, $R^2=0.348$, Adjusted $R^2=0.251$, $F=3.581$, $p=0.004$. VAS: Visual Analog Scale; SAI: State Anxiety Inventory; VIF: variance inflation factor. Bold values indicate statistically significant results ($p<0.05$).

pain, and comfort levels. In addition, collecting the data during the pre-discharge period may have resulted in responses being influenced by temporary emotional states, as women were likely still experiencing the physical and emotional effects of the procedure. Furthermore, the inclusion of women with specific sociodemographic and clinical characteristics may limit the generalizability of the findings to other cultural and clinical populations.

Another limitation was the inability to control for potential confounders, such as the amount of pre-procedural information, previous pregnancy loss, and individual coping mechanisms. Participants' emotional distress related to pregnancy loss may also have influenced pain perception and self-reported pain scores. In addition, including women with fetal anomalies and intrauterine fetal death in the same sample may have introduced heterogeneity because these conditions involve different psychological and clinical processes. Finally, comfort was assessed using a scale developed for post-anesthesia patients, which may not have fully captured comfort specific to medical termination.

## CONCLUSION

This study demonstrated that women undergoing medical termination experienced anxiety both before and after the procedure, severe pain during the procedure, and moderate pain after the procedure. Higher anxiety levels were associated with lower post-procedural comfort. Therefore, routine assessment of anxiety before uterine evacuation and the implementation of targeted interventions to reduce anxiety may help improve women's procedural comfort. Although effective pain management remains essential, care should address both the physical and psychological dimensions of the medical termination experience.

Accordingly, women's health nurses should play an active role in pain management, provide individualized pre-procedural counseling to reduce fear and anxiety, and deliver evidence-based, holistic care tailored to each woman's physical and psychological needs throughout the medical termination process.

## ETHICAL ASSESSMENT

Ethics committee approval was obtained from Çukurova University's Faculty of Medicine Ethics Committee (No. 135-41; July 14, 2023).

## AUTHORS' CONTRIBUTIONS

**SKY:** Conceptualization, Formal Analysis, Methodology, Software, Supervision, Visualization, Writing – original draft, Writing – review & editing. **MNS:** Conceptualization, Data curation, Methodology, Writing – original draft, Writing – review & editing. **FIA:** Investigation, Resources, Validation. **EN:** Investigation, Validation.

## DATA AVAILABILITY STATEMENT

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

## REFERENCES


1. Uncu B, Doğan E. Epigenetics and pregnancy. *İzmir Katip Çelebi* Univ J Fac Health Sci. *2024;9*(1):161-5. https://doi.org/10.61399/ikcusbfd.1310676
2. Ağaoğlu Z, Tanacan A, Haksever M, Coşkun H, İpek G, Denizli R, et al. Retrospective analysis of the indications, methods, and complications of pregnancy termination. Turk J Obstet Gynecol. 2024;21(4):273-9. https://doi.org/10.4274/tjod.galenos.2024.88886
3. Kirtis E, Alkan Bulbul G, Kandemir H, Metin T, Sanhal CY, Mendilcioglu İİ. Late termination of pregnancy: a retrospective evaluation of 243 pregnancies ın a tertiary medical center. Akdeniz Med J. 2025;11(2):157-63. https://doi.org/10.53394/akd.1345750
4. Choi H, Kim HS, Sakong J. Rates and subsequent clinical course of fetal congenital anomalies detected by prenatal targeted ultrasonography: a retrospective observational study. *J Yeungnam Med Sci. 2023;40*(3):268-75. https://doi.org/10.12701/jyms.2022.00514
5. Carmen Prodan N, Hoopmann M, Jonaityte G, Kagan KO. How to do a second trimester anomaly scan. *Arch Gynecol Obstetr. 2023;307*(4):1285-90. https://doi.org/10.1007/s00404-022-06569-2
6. Gök K, Özgül A, Yılmaz E, Köse O, Bostancı MS, Özden S. Retrospective analysis of pregnancy termination cases: evaluation of 177 cases. *Sakarya Med J. 2022;12*(1):116-21. https://doi.org/10.31832/smj.1061707
7. Turkyilmaz G, Atan YS, Erturk E, Turkyilmaz S. *Psychological reactions in women with pregnancy termination due to fetal anomaly. Gynecol Obstetri Reprod Med.* 2021;27(3):202-6. https://doi.org/10.21613/GORM.2021.1145
8. Irmscher L, Marx R, Linke M, Zimmermann A, Drössler S, Berth H. Anxiety, depression, somatization and psychological distress before and 2–6 years after a late termination of pregnancy due to fetal anomalies. BMC Women's Health. 2024;24(1):255. https://doi.org/10.1186/s12905-024-03082-3
9. Kaydırak MM, Aslan E. The medical termination of pregnancy: an example nursing support programme. *J Anatolia Nurs Health Sci. 2020;23*(3):430-43. https://doi.org/10.17049/ataunihem.480328
10. Jackson E, Kapp N. Pain management for medical and surgical termination of pregnancy between 13 and 24 weeks of gestation: a systematic review. *BJOG Int J Obstetr Gynaecol. 2020;127*(11):1348-57. https://doi.org/10.1111/1471-0528.16212

11. Eti Aslan F. The assessment methods of pain. J Cumhuriyet Univ Sch Nurs. 2002;6(1):9-16.
12. Öner N, Compte A. State-trait anxiety inventory, Handbook. 2. Istanbul: Print Boğaziçi University Publications; 1995
13. Üstündağ H, Eti Aslan F. Adaptation of the perianesthesia comfort scale into turkish. *Türkiye Klinikleri J Nurs Sci. 2010;2*(2):94-9.
14. Bhat A, Byatt N. Infertility and perinatal loss: when the bough breaks. *Curr Psychiatry Rep. 2016;18*(3):1-11. https://doi.org/10.1007/s11920-016-0663-8
15. Fockler ME, Ladhani NNN, Watson J, Barrett JFR. Pregnancy subsequent to stillbirth: medical and psychosocial aspects of care. *Semin Fetal Neonatal Med. 2017;22*(3):186-92. https://doi.org/10.1016/j.siny.2017.02.004
16. Opee J, Mayanja SB, Kayondo M, Tibaijuka L, Bongomin F, Orach CG, et al. Proportion and factors associated with intra-procedural pain among women undergoing manual vacuum aspiration for incomplete abortion at Mbarara Regional Referral Hospital Uganda. Pan Afr Med J. 2024;49:63. https://doi.org/10.11604/pamj.2024.49.63.39955
17. Güçlü O, Şenormancı G, Tüten A, Gök K, Şenormancı Ö. Perinatal grief and related factors after termination of pregnancy for fetal anomaly: one-year follow-up study. Arch Neuropsychiatry. *2021;58*(3):221-7. https://doi.org/10.29399/npa.25110
18. Farren J, Jalmbrant M, Ameye L, Joash K, Mitchell-Jones N, Tapp S, et al. Post-traumatic stress, anxiety and depression following miscarriage or ectopic pregnancy: a prospective cohort study. *BMJ Open. 2016;6*(11):e011864. https://doi.org/10.1136/bmjopen-2016-011864
19. Kabalcıoğlu Bucak F, Toker E, Asoğlu M. Women who apply therapeutic abortus' state and trait anxiety levels and related factors. J Health Acad. *2018;5*(4):256-63. https://doi.org/10.5455/sad.13-1543399784
20. Davoudian T, Gibbins K, Cirino NH. Perinatal loss: the impact on maternal mental health. *Obstetr Gynecol Surv. 2021;76*(4):223-33. https://doi.org/10.1097/OGX.0000000000000874
21. Aydingülü N, Arslan S. Comfort levels of patients at early postoperative period. J Fac Med Kahramanmaraş Sütçü İmam Univ. *2021;16*(3):401-6. https://doi.org/10.17517/ksutfd.811420
22. Li XR, Zhang WH, Williams JP, Li T, Yuan JH, Du Y, et al. A multicenter survey of perioperative anxiety in China: pre- and postoperative associations. *J Psychosomatic Res. 2021;147:110528*. https://doi.org/10.1016/j.jpsychores.2021.110528

https://doi.org/10.1590/1806-9282.20251775-ER

# ERRATUM: Clinical profile of acute myocardial infarction: the Registry of Acute Medical Emergencies registry

In the manuscript: Silva PGM de B e, Soeiro AM, Balada R, Macedo L, Lacerda MDV de, Correia LVA, et al. Clinical profile of acute myocardial infarction: the Registry of Acute Medical Emergencies registry. Rev Assoc Med Bras. 2026;72(6):e20251775. https://doi.org/10.1590/1806-9282.20251775, on page 1:

**Where it reads:**
Sofia Gabriela Gregolini Catellani

**It should read:**
Sofia Gabriella Gregolini Catellani

Roseli Nomura
*Editor-in-Chief*

CC BY