JOURNAL OF MEDICINE
2020
FORBES

FORBES JOURNAL OF MEDICINE

Volume 7 | Year 2026



www.forbestip.org
e-ISSN: 2757-5241

galenos
Publishing House

FORBES
JOURNAL OF MEDICINE

e-ISSN: 2757-5241

Year 2026
Volume: 7

Forbes Journal of Medicine is the publication of İzmir Buca Seyfi Demirsoy Training and Research Hospital.

FJM is an open access, free, peer-reviewed journal that operates under a continuous publication model.

Please refer to the journal's webpage **(https://forbestip.org)** for "Aims and Scope", "Ethical Policy", "Checklist for Authors" and "Manuscript Preparation".

The editorial and publication process of Forbes Journal of Medicine are shaped in accordance with the guidelines of the ICMJE, WAME, CSE, COPE, EASE, and NISO.

Forbes Journal of Medicine is indexed in **Ulakbim TR Dizin, Turkey Citation Index, EmCare, EmBiology, EMBASE, EBSCO, Türk Medline, J-Gate, Gale** and **DOAJ**.

The journal is published electronically.

Owner: M. Yekta Öncel, M.D. Prof. on Behalf of Buca Seyfi Demirsoy Training and Research Hospital Head Physician

Editor in Chief: M. Yekta Öncel, M.D. Prof.

Administrative Office

Chief Physician's Office, Buca Seyfi Demirsoy Training and Research Hospital, Buca/İzmir, Türkiye

Tel: +90 (232) 452 52 52 /4293 **E-mail:** forbestipdergisi@gmail.com

Publication Type

Scholarly Periodical

Publisher

GALENOS PUBLISHING HOUSE

Molla Gürani Neighborhood, Kaçamak Street No: 21/1, 34093 Fındıkzade, Fatih, Istanbul, Türkiye

Tel: +90 (530) 177 30 97 **E-mail:** info@galenos.com.tr/yayin@galenos.com.tr **Web:** www.galenos.com.tr

Publisher Certificate No:: 14521

Online Publication Date

Year 2026

FORBES
JOURNAL OF MEDICINE

Editorial Board

Founder

M. Yekta Öncel, M.D. Prof.
Buca Seyfi Demirsoy Training and Research Hospital/Head Physician/İzmir, Türkiye

Editor in Chief

M. Yekta Öncel, M.D. Prof.
İzmir Kâtip Çelebi University Faculty of Medicine/Neonatology/İzmir, Türkiye
E-mail: mehmetyekta.oncel@ikcu.edu.tr

Editors

Banu İşbilen Başok, M.D. Prof.
University of Health Sciences Türkiye, İzmir Faculty of Medicine/Medical Biochemistry/İzmir, Türkiye
E-mail: banu.basok@saglik.gov.tr

Berna Dirim Mete, M.D. Prof.
İzmir Democracy University Faculty of Medicine/Radiology/İzmir, Türkiye
E-mail: berna.dirim@idu.edu.tr

Deniz Çankaya, M.D. Prof.
University of Health Sciences Türkiye, Gülhane Training and Research Hospital/Orthopedics and Traumatology/Ankara, Türkiye
E-mail: deniz.cankaya@sbu.edu.tr

Hanifi Soylu, M.D. Prof.
Canada University of Manitoba/Clinical Pharmacology-Clinical Epidemiology/Canada
E-mail: hasoylu@hotmail.com

Umut Varol, M.D. Prof.
İzmir Democracy University Faculty of Medicine/Medical Oncology/İzmir, Türkiye
E-mail: umut.varol@idu.edu.tr

Hakan Gülmez, M.D. Assoc. Prof. (Responsible Manager)
İzmir Democracy University Faculty of Medicine/Family Medicine/İzmir, Türkiye
E-mail: hakan.gulmez@idu.edu.tr

Suzan Şahin, M.D. Assoc. Prof.
İzmir Democracy University Faculty of Medicine/Neonatology/İzmir, Türkiye
E-mail: suzan.sahin@idu.edu.tr

Elif Keleş, MD, PhD
Northwestern University Feinberg School of Medicine/Radiology Hybrid and Machine Intelligence Lab / Chicago, Illinois, USA
E-mail: elif.keles@northwestern.edu

Linguistic Editor

Ahu Pakdemirli, M.D. Assoc. Prof.
University of Health Sciences Türkiye, İzmir Faculty of Medicine/Physiology/İzmir, Türkiye
E-mail: ahu.pakdemirli@sbu.edu.tr
ORCID ID: 0000-0001-9224-3007

Biostatistics Editor

Ferhan Elmalı, Ph.D. Prof.
İzmir Kâtip Çelebi University Faculty of Medicine/Biostatistics/İzmir, Türkiye
E-mail: ferhan.elmalı@ikcu.edu.tr
ORCID ID: 0000-0002-1967-1811

Publishing Secretary

Gül Aslan
Buca Seyfi Demirsoy Training and Research Hospital/R&D Department/İzmir, Türkiye
E-mail: gul.aslan2@saglik.gov.tr

Editorial Advisory Board

Duygu Adiyaman, M.D.
Universitätsklinikum Ulm Klinik für Frauenheilkunde und Geburtshilfe Pränatale Medizin/Perinanology/Germany

Servet Akar, M.D.
İzmir Kâtip Çelebi University Faculty of Medicine/Rheumatology/İzmir, Türkiye

Karel Allegaert, Ph.D.
Katholieke Universiteit/Development and Regeneration/Leuven, Belgium Erasmus Medical Center/Clinical Pharmacy/Rotterdam, Holland

Vefik Arıca, M.D.
Yalova University Faculty of Medicine/Child Health and Diseases/Yalova, Türkiye

Ahmad Bassiouny, M.D.
Alexandria University Faculty of Science/Medical Biochemistry/Egypt

Hafize Öztürk Can, Ph.D.
Ege University Faculty of Health Sciences/Department of Midwifery/İzmir, Türkiye

Fuat Emre Canpolat, M.D.
Ankara City Hospital/Neonatology/Ankara, Türkiye

Gönül Çatlı, M.D.
İstinye University, İstanbul Liv Hospital/Pediatric Endocrinology/İstanbul, Türkiye

Mehmet Nevzat Çizmeci, M.D.
The Hospital for Sick Children, University of Toronto/Neonatology/Canada

Esra Duğral, M.D.
Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital/Physiology/İzmir, Türkiye

Bumin Nuri Dündar, M.D.
İzmir Kâtip Çelebi University Faculty of Medicine/Pediatric Endocrinology/İzmir, Türkiye

Umut Elboğa, M.D.
Gaziantep University Faculty of Medicine/Nuclear Medicine/Gaziantep, Türkiye

Yasser Elsayed, M.D.
University of Canada Manitoba/Neonatology/Canada

Ömer Erdeve, M.D.
Ankara University Faculty of Medicine/Neonatology/Ankara, Türkiye

Atilla Ersen, M.D.
Buca Seyfi Demirsoy Training and Research Hospital/Child Neurology/İzmir, Türkiye

Ali Guliyev Arif, M.D.
Medera Hospital/Child Health and Diseases/Baku, Azerbaijan

Eda Karadağ Öncel, M.D.
Dokuz Eylül University Faculty of Medicine/Pediatric Infectious Diseases/İzmir, Türkiye

Soner Sertan Kara, M.D.
Aydın Adnan Menderes University Faculty of Medicine/Pediatric Infectious Diseases/Aydın, Türkiye

Konstantin Kenigsberg, M.D.
MRI-Center OrthoClinic/Radiology/Belarus

Ramazan Özdemir, M.D.
İnönü University Faculty of Medicine/Neonatology/Malatya, Türkiye

M. Burak Öztop, M.D.
Ministry of Health General Directorate of Health Promotion/General Surgery/Ankara, Türkiye

Betül Taşpınar, Ph.D.
İzmir Democracy University Faculty of Health Sciences/Physiotherapy and Rehabilitation/İzmir, Türkiye

Murat Tutanç, M.D.
University of Health Sciences Türkiye, Bursa Faculty of Medicine/Child Health and Diseases/Bursa, Türkiye

Semra Ulusoy Kaymak, M.D
Ministry of Health General Directorate of Health Services/Mental Health and Diseases/Ankara, Türkiye

Sevil Üzer, Ph.D.
Academic Teaching Hospital of the University of Heidelberg/Physical Medicine and Rehabilitation/Maulbronn, Germany

İlhan Yaylım, M.D.
İstanbul University Aziz Sancar Experimental Medicine Research Institute/Molecular Medicine/İstanbul, Türkiye

Sadık Yurttutan, M.D.
Kahramanmaraş Sütçü İmam University/Neonatology/Kahramanmaraş, Türkiye

Natalia Zarbailov, M.D.
State University of Medicine and Pharmacy/Family Medicine/Moldova

FORBES
JOURNAL OF MEDICINE

Editorial

Dear Readers,

As the editorial board of Forbes Journal of Medicine, we are delighted to present the 2026 editorial of our journal with increasing excitement and motivation after our sixth year of publication. In the past year, a total of 39 scientific studies were published in our journal in the fields of Medicine and Health Sciences, including 3 review articles, 34 original research articles, and 2 case reports. As a result of the publication evaluation and printing processes meticulously carried out by the Forbes Journal of Medicine Editorial Board since the first issue, our journal continues to be indexed in our national index ULAKBİM (TR Dizin), Turkish Citation Index, and Turkish Medline, which have a very important place in academic publishing, as well as in international academic indexes such as EBSCO, J-Gate, GALE, and DOAJ. In addition, as of the past year, our journal has also taken its place in the prestigious international indexes Embase, EmCare, and EmBiology. We would like to share that, in line with our goal of further strengthening our journal and achieving success, especially internationally, we are continuing our application process to important international indexes, primarily ESCI and SCOPUS.

As Forbes Journal of Medicine, we closely follow current developments in scientific publication processes, and in this regard, we would like to share that as of the 2026 publication year, our journal will be published only in English and will have transitioned to a continuous publication model. We would like to express our sincere gratitude to the authors who submitted their work to our journal, to the editors and advisory board members who evaluated them, and to our reviewers.

Sincerely,

Professor M. Yekta ÖNCEL, MD
Editor in Chief
Forbes Journal of Medicine

DOI: 10.4274/forbes.galenos.2026.94557
Forbes J Med 2026;7:1-6

Original Article

Prognostic Significance of Immune-Inflammation Markers, Lodds and LNR in Locally Advanced Gastric Cancer

Melis KILINÇ[1], Mehmet UZUN[2], İlkay Tuğba ÜNEK[3]

[1]Dokuz Eylül University Faculty of Medicine, Department of Internal Medicine, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Division of Medical Oncology, Clinic of Internal Medicine, İzmir, Türkiye
[3]Dokuz Eylül University Faculty of Medicine, Division of Medical Oncology, Department of Internal Medicine, İzmir, Türkiye

Cite as: Kılınç M, Uzun M, Ünek İT. Prognostic significance of immune-inflammation markers, lodds and LNR in locally advanced gastric cancer. Forbes J Med. 2026;7:1-6

ABSTRACT

Objective: Gastric cancer continues to be a pressing issue in global health. This research seeks to examine the relationship between immune system-related inflammatory markers.

Methods: The research involved 184 cases of locally advanced gastric cancer diagnosed between January 2010 and January 2021. In light of its retrospective methodology, the study did not necessitate informed consent, as per institutional ethical guidelines. Receiver operating characteristic analysis was applied to establish the optimal threshold values for the systemic inflammatory response index (SIRI), systemic inflammatory index (SII), prognostic nutritional index (PNI), pan-immune inflammation value (PIV), neutrophil-lymphocyte ratio, lymph node ratio (LNR) and log odds of positive lymph nodes (LODDS). In these groups, survival outcomes were analyzed using the Kaplan-Meier method. The association between mortality and risk factors was assessed using Cox regression analysis. All tests were deemed statistically significant if the $p<0.05$.

Results: The SIRI, SII, PIV, PNI, LNR and LODDS values were shown to be correlated with overall survival (OS) duration ($p<0.05$). Cox regression model with multiple variables identified that PNI was an independent determinant of OS ($p=0.045$).

Conclusion: The conclusions drawn from this research suggest that immune-inflammation markers, along with the LNR and LODDS values of patients with local advanced stage gastric cancer diagnosis may be used as prognostic factors in routine clinical practice. Detection of these immune-inflammation markers, LNR and LODDS values may guide clinicians in prognostic evaluation as well as the creation of personalized treatment approaches.

Keywords: Gastric cancer, LNR, LODDS

Received: 18.01.2026
Accepted: 25.02.2026
Epub: 25.02.2026
Publication Date: 09.03.2026

Corresponding Author:
Melis KILINÇ
Dokuz Eylül University Faculty of Medicine, Department of Internal Medicine, İzmir, Türkiye
drmeliskilinc@gmail.com
ORCID ID: 0000-0002-4288-9632

INTRODUCTION

Gastric cancer has the 6th-highest incidence after breast, prostate, lung, colorectal, and cervical cancer, according to GLOBOCAN 2022 predictions for global cancer statistics. With respect to mortality, it ranks 7th, with a rate of 6.1% after lung, breast, colorectal, liver, prostate, and cervical cancers.[1] Accumulating evidence indicates that inflammation plays a crucial role in cancer initiation and progression.[2] In recent years, immune-inflammation-based markers for patients with cancer have been developed using laboratory parameters that are readily available in routine clinical practice, and their prognostic significance has been investigated. The systemic inflammatory response index (SIRI), the systemic inflammatory index (SII), the pan-immune inflammation value (PIV), the prognostic nutritional index (PNI), and the neutrophil-lymphocyte ratio (NLR) may be included among these biomarkers.[3-6]

Pathological lymph node (pN) classification is determined by the number of lymph nodes removed during gastrectomy that are found to be metastatic on pathological examination. The need to develop different classification systems has emerged with the aim of increasing the reliability of lymph node staging.[7] The percentage of metastatic lymph nodes relative to the total number of excised lymph nodes,

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

referred to as the lymph node ratio (LNR), is calculated after surgical resection for locally advanced gastric cancer and serves as a marker of poor prognosis in high-risk patients.[8-10] The LNR classification system has deficiencies: its prognostic estimation power has been reported to decrease when LNR is 0 or 1. The classification determined using the log odds of positive lymph nodes (LODDS), known as the LODDS value, was identified as better at determining disease prognosis than pN or LNR classifications in gastric cancer. The LODDS value is calculated as the logarithm of the ratio of (metastatic lymph nodes +0.5) to (negative lymph nodes +0.5), and is a novel prognostic factor.[11-15]

In this study, correlations between the biomarkers SIRI, PIV, NLR, SII, PNI, LNR, and LODDS and prognosis were investigated in patients with locally advanced gastric cancer.

METHODS

Patient Population

From January 2010 to January 2021, the study included 184 patients who were diagnosed with locally advanced-stage gastric cancer and had tests and treatment planned at Dokuz Eylül University, Oncology Clinic; they were aged 18 years or older and had no history of secondary solid or hematological tumors. Our study had a retrospective design, and the clinical information, laboratory and pathology results, and treatments administered to patients were recorded. Overall survival (OS) was defined as the interval from the surgical procedure to death or last follow-up. Locally advanced-stage gastric cancer was defined as cT1b-T4 disease in patients without clinically detectable distant metastasis, with or without lymph node involvement.[16]

Immune-Inflammation Markers

Scores for immune-inflammation-based markers were calculated in accordance with definitions and calculation methods reported in the literature, using patients' hematological and biochemical parameters obtained before treatment. The following formulas were used:

NLR: neutrophil (10^3/μL)/lymphocytes (10^3/μL)

SIRI: [neutrophil (10^3/μL) x monocytes (10^3/μL)]/ lymphocytes (10^3/μL)

PIV: [neutrophil (10^3/μL) x monocytes (10^3/μL) x platelets (10^3/μL)]/lymphocytes (10^3/μL)

SII: [neutrophil (10^3/μL) x platelet (10^3/μL)]/lymphocyte (10^3/μL)

PNI: [albumin(g/dL) x 10] + [lymphocytes (10^3/μL) x 0.005]

To determine cut-off values for SIRI, PIV, NLR, SII, PNI, LNR, and LODDS, receiver operating characteristic (ROC) curve analysis was used, and OS analysis was performed. The analysis results in accordance with the ROC curve determined the mean cut-off values were 1.45 [area under the curve (AUC): 0.574 (95% confidence interval (CI): 0.50-0.64, p=0.04] for SIRI, 391.9 [AUC: 0.567 (95% CI: 0.49-0.63, p=0.06)] for PIV, 2.4 [AUC: 0.603 (95% CI: 0.53-0.67, p=0.004)] for NLR, 637.5 [AUC: 0.593 (95% CI: 0.52-0.66, p=0.01)] for SII, 48 [AUC: 0.606 (95% CI: 0.53-0.67, p=0.003)] for PNI, 0.28 [AUC: 0.731 (95% CI: 0.66-0.79, p<0.001)] for LNR and 0.40 [AUC: 0.740 (95% CI: 0.67-0.80, p<0.001)] for LODDS.

LNR and LODDS Classification

LNR is determined by dividing the number of metastatic lymph nodes by the total number of lymph nodes removed. In addition to cases where LNR is 0 or 1, the LNR value is divided into 5 groups in intervals of 0.1. The LNR classification is LNR 1(LNR =0), LNR 2 (0< LNR ≤0.10), LNR 3 (0.1< LNR ≤0.2), LNR 4 (0.2< LNR ≤0.5) and LNR 5 (LNR >0.5). The LODDS value is determined by calculating the logarithm of the ratio of (metastatic lymph nodes +0.5) to (negative lymph nodes +0.5). The categorization of LODDS classes uses intervals of 0.5 and follows a structure comparable to that of LNR classes. The LODDS classification is LODDS 1(LODDS ≤-1.5), LODDS 2 (<-1.5 LODDS ≤-1), LODDS 3 (<-1 LODDS ≤-0.5), and LODDS 4 (-0.5< LODDS ≤0).

Statistical Analysis

Data were analyzed using SPSS 22.0. After descriptive analysis, the normality of quantitative variables was assessed using the Shapiro-Wilk and Kolmogorov-Smirnov tests. Comparisons of quantitative variables that did not follow a normal distribution between the mortality and progression groups were performed using the Mann-Whitney U test for independent samples. Comparisons of quantitative variables with a bell-shaped distribution were performed using the Student's t-test. Results are reported as median (range) and mean (standard deviation). Comparisons of categorical variables were performed using Fisher's exact test and the chi-square test. Data are reported as numbers (percentages).

Predictive values of immune-inflammatory markers for mortality and progression/recurrence were determined by analysis of the ROC curve. AUC measurements were determined using the Youden index (YI) for parameters with p value <0.05. For each parameter, the value with the highest YI was determined as the cut-off, and the true positive and true negative rates for these cut-offs were calculated. According to these threshold values, patients

were divided into low- and high-risk groups. Survival analysis for these groups was performed using the Kaplan-Meier method. The log-rank test was used to compare median OS values between different risk groups for each parameter. Identification of factors influencing mortality or progression was performed using Cox regression analysis. Findings are reported with 95% confidence. For all tests, $p<0.05$ was considered statistically significant.

Ethics Committee Approval

The study was completed after receiving approval from the Dokuz Eylül University Non-Interventional Research Ethics Committee (dated 10.05.2023 and numbered 8059-GOA).

RESULTS

Among patients diagnosed with locally advanced gastric cancer, 118 male and 66 female patients were included. The median age of the 184 patients was 61 years (range 23–85). According to the Eastern Cooperative Oncology Group (ECOG) performance classification, 48.9% of patients were categorized as ECOG 0, 38% as ECOG 1, 11.4% as ECOG 2, and 1.6% as ECOG >2. Regarding treatment, 52.2% of patients had neoadjuvant chemotherapy, 91.8% had surgical treatment, and 63% had adjuvant radiotherapy. When pathological subtypes were evaluated, the most frequent subtype was adenocarcinoma (77.2%), followed by signet ring cell carcinoma (15.2%). Table 1 provides the clinical and pathological details of the study cohort; Table 2 provides the laboratory parameters used to calculate the immune-inflammation-based markers.

Table 1. Patient clinical and pathological features

Age median, years (min-max)	**61.0 (23-85)**
Sex, n (%)	
Man	118.0 (64.1)
Woman	66.0 (35.9)
Performance status, n (%)	
ECOG 0	90.0 (48.9)
ECOG 1	70.0 (38.0)
ECOG 2	21.0 (11.4)
ECOG >2	3.0 (1.6)
Comorbid disease, n (%)	
HT	47.0 (25.5)
DM	26.0 (14.1)
CAD	4.0 (2.2)
COPD	18.0 (9.8)
Operation, n (%)	169.0 (91.8)
LNR, n (%)	
LNR 1 (LNR =0)	30.0 (17.9)
LNR 2 (0< LNR ≤0.10)	50.0 (29.8)
LNR 3 (0.1< LNR ≤0.2)	24.0 (14.3)
LNR 4 (0.2< LNR ≤0.5)	45.0 (26.8)
LNR 5 (LNR >0.5)	19.0 (11.3)
LODDS classification, n (%)	
LODDS 1 (LODDS ≤-1.5)	32.0 (18.9)
LODDS 2 (-1.5< LODDS ≤-1)	35.0 (20.7)
LODDS 3 (-1< LODDS ≤-0.5)	43.0 (25.4)
LODDS 4 (-0.5< LODDS ≤0)	59.0 (34.9)
Adjuvant RT, n (%)	116.0 (63)
Neoadjuvant CT, n (%)	96.0 (52.2)
Tumor histopathology, n (%)	
Adenocarcinoma	142.0 (77.2)
Signet ring cell carcinoma	28.0 (15.2)
Other	14.0 (7.6)
LNR, median (min-max)	0.1 (0-1)
LODDS, median (min-max)	-0.8 (-2.1-1.2)

ECOG: Eastern Cooperative Oncology Group, HT: Hypertension, DM: Diabetes mellitus, CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, LNR: Lymph node ratio, LODDS: Log probability of positive lymph nodes, RT: Radiotherapy, CT: Chemotherapy, min: Minimum, max: Maximum

Table 2. Immune-inflammation markers of patients

Parameter	**Median**	**Min-max**
NEU	4.5	0.9-24.5
LYM	1.8	0.4-4.1
HB	11.6	6.3-16.3
PLT	250.0	99.0-587.0
Albumin	3.8	1.8-4.8
SIRI	1.4	0.1-26.1
PIV	391.9	8.3-5472
NLR	2.4	0.7-36.5
SII	637.5	82.6-8788
PNI	48.0	21.5-59.5

NEU: Neutrophil, LYM: Lymphocyte, HB: Hemoglobin, PLT: Platelets, SIRI: Systemic inflammatory response index, PIV: Pan-immune inflammation value, NLR: Neutrophil lymphocyte ratio, SII: Systemic inflammatory index, PNI: Prognostic nutritional index, min: Minimum, max: Maximum

The median OS of patients was identified as 35.5±4.6 months. Patients with age under or equal to 61 years had median OS of 44.9±7.8 months, while patients over 61 years of age had median OS of 28.4±4.2 months and a notable statistical variation was detected across the age categories ($p=0.020$).

For patients with SIRI value >1.45, the OS was 44.9±19.6 months, while patients with SIRI ≤1.45 had OS of 27.9±3.2 months and the survival durations between the groups were found to be notable statistical variation ($p=0.009$). Patients with PIV ≤391.9 had OS of 44±8 months, while patients with PIV >391.9 had OS of 28.1±3.2 months and there a notable statistical variation for OS durations across

the PIV risk categories (p=0.048). For cases with NLR >2.4, the OS was 44.0±7.1 months, while cases with NLR ≤2.4 had OS of 29.2±3.5 months and there was no notable statistical variation across the groups (p=0.05). Patients with SII ≤637.5 had median OS of 44.9±8.7 months, while patients with SII >637.5 had median OS of 28.4±2.6 months (p=0.031). Patients with PNI >48 had median OS that was statistically significantly longer compared to patients with PNI ≤48 (44.9±10.0 months and 28.0±4.5 months, respectively, p=0.017). Cases having LNR >0.28 had survival durations that were statistically significantly shorter compared to those with LNR ≤0.28 (p<0.001). For cases having LODDS >-0.40, the survival durations were statistically significantly worse than for those having LODDS ≤-0.40 (p<0.001) (Table 3). According to the results of a multivariable Cox regression model developed using immune-inflammation markers PNI values were independent determinants of OS (p=0.045) (Table 4).

DISCUSSION

In this study, significant correlations were observed between immune-inflammation-based markers (SIRI, PIV, SII, PNI), LNR, and LODDS and OS in patients with locally advanced gastric cancer.

SIRI is an index derived from levels of neutrophils, monocytes, and lymphocytes, reflecting the relationship between inflammatory processes and immune function.

In our study, those with SIRI values >1.45 had a significantly shorter median OS. This finding is consistent with the literature reporting that elevated SIRI is associated with an unfavorable prognosis in gastric cancer. Ren et al.[17] similarly associated higher SIRI values with poor OS; however, they did not find an optimal cut-off value. Another study proved that the SIRI index was an independent determinant in patients with gastric cancer who underwent radical gastrectomy.[18] In the literature, the five-year survival rate for gastric cancer cases with high preoperative NLR has been reported to be considerably shorter than that for

Table 3. Correlation of immune-inflammation markers, LNR and LODDS with overall survival

	Kaplan-meier analysis			Cox regression analysis (univariate)	
Parameter	**Median OS (month)**	**95% CI**	**p**	**HR**	**p**
Age ≤61 years >61 years	 44.9±7.8 28.4±4.2	 29.6-60.1 20.1-36.7	**0.019**	1.6 (1.1-2.3)	**0.020**
Sex Man Woman	 42.8±5.9 33.0±3.3	 31.3-54.3 26.6-39.5	0.868	1.0 (0.7-1.5)	0.868
SIRI ≤1.45 >1.45	 44.9±19.6 27.9±3.2	 6.4-83.3 21.6-34.3	**0.009**	1.7 (1.1-2.5)	**0.010**
PIV ≤391.9 >391.9	 44.0±8.0 28.1±3.2	 28.4-59.6 21.8-34.5	**0.048**	1.5 (1.0-2.2)	0.050
NLR ≤2.4 >2.4	 44.0±7.1 29.2±3.5	 30.1-57.9 22.3-36.2	0.050	1.5 (1.0-2.2)	0.051
SII ≤637.5 >637.5	 44.9±8.7 28.4±2.6	 27.9-61.9 23.4-33.5	**0.031**	1.5 (1.0-2.2)	**0.032**
PNI >48.0 ≤48.0	 44.9±10.0 28.0±4.5	 25.3-64.4 19.1-36.7	**0.017**	1.6 (1.1-2.3)	**0.019**
LNR ≤0.28 >0.28	 103.5±--- 24.4±3.3	 ----- 17.9-30.8	**<0.001**	3.7 (2.4-5.6)	**<0.001**
LODDS ≤-0.40 >-0.40	 103.5±--- 24.4±3.3	 ----- 17.9-30.8	**<0.001**	3.7 (2.4-5.6)	**<0.001**

SIRI: Systemic inflammatory response index, PIV: Pan-immune inflammation value, NLR: Neutrophil lymphocyte ratio, SII: Systemic inflammatory index, PNI: Prognostic nutritional index, LNR: Lymph node ratio, LODDS: Log probability of positive lymph nodes, CI: Confidence interval, OS: Overall survival, HR: Hazard ratio

Table 4. Cox regression model with multiple variables

	Cox regression model with multiple variables		
	HR	**95% CI**	**p**
SIRI ≤1.45 (92.0) vs. >1.45 (92.0)	2.2	1.0-5.2	0.050
PIV ≤391.9 (92.0) vs. >391.9 (92.0)	0.7	0.3-1.8	0.481
NLR ≤2.4 (92.0) vs. >2.4 (92.0)	0.7	0.3-1.5	0.383
SII ≤637.5 (92.0) vs. >637.5 (92.0)	1.4	0.6-2.8	0.421
PNI >48.0 (92.0) vs. ≤48.0 (91.0)	1.5	1.0-2.3	**0.045**

SIRI: Systemic inflammatory response index, PIV: Pan-immune inflammation value, NLR: Neutrophil lymphocyte ratio, SII: Systemic inflammatory index, PNI: Prognostic nutritional index, CI: Confidence interval, HR: Hazard ratio

cases with low NLR.[19,20] According to the results of this study, no meaningful relationship was found between NLR and survival (Table 2, p=0.05).

With locally advanced gastric cancer, patients with SII >637.5 had shorter median OS duration compared to patients with SII ≤637.5. In the literature, cases with elevated SII exhibited shorter OS.[21] A study including a broad group of patients found that increased SII prior to surgery is an independent determinant of adverse prognosis in gastric cancer.[20]

In cases with locally advanced gastric cancer, cases with PNI >48 had statistically significantly longer median OS compared to patients with PNI ≤48. In our study, in light of the results from the Cox regression model with multiple variables developed using immune-inflammation markers, cases with a PNI ≤48 demonstrated a substantial increase in mortality (p=0.045). In the literature, increased PNI was associated with longer OS. In these studies, multivariate analyses showed that a low PNI value was an independent determinant of worse OS.[22,23] Identifying PNI as an independent risk factor for OS reiterates the importance of nutritional status for cancer prognosis. Providing nutritional support to patients with gastric cancer in the preoperative period has the potential to improve survival. Additionally, combining PNI with other inflammation markers may be beneficial in creating stronger prognostic models.

In our study, the LNR and LODDS values, which evaluate lymph node metastasis, were correlated with prognosis. In patients having an LNR >0.28, median OS was statistically significantly shorter in comparison to those with an LNR ≤0.28. In the literature, the percentage of lymph nodes involved by metastasis has served as a prognostic indicator, independent of the number of lymph nodes removed during the operation. An increase in this proportion was associated with decreased OS.[24,25] As part of our research, cases with LODDS >-0.40 were found to have statistically significantly shorter median OS relative to cases with LODDS ≤-0.40. Determining pN stages based solely on the lymph node count with metastasis and the classification variation between <15 and ≥15 lymph nodes highlighted the need for stronger prognostic measures, such as LNR and LODDS, to predict outcomes. LODDS accounts for both metastatic and non-metastatic lymph node counts. It was developed to provide an accurate prognostic assessment for patients with pN0 and <15 removed lymph nodes. In the literature, as the LODDS degree increases, prognosis is affected, and LODDS appears to be a more reliable prognostic indicator than pN.[26]

CONCLUSION

Our study demonstrated the usefulness of the immune-inflammation markers LNR and LODDS as prognostic factors in clinical practice for patients with locally advanced gastric cancer. Due to the ease of measurement using routine blood tests and the number of lymph nodes with and without metastasis in pathology reports, they offer practical and economic benefits. Detection of these immune-inflammation markers, LNR and LODDS values, may guide clinicians in predicting patient prognosis and planning optimal treatment.

Ethics

Ethics Committee Approval: The study was completed after receiving approval from the Dokuz Eylül University Non-Interventional Research Ethics Committee (dated 10.05.2023 and numbered 8059-GOA).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.K., Concept: M.U., İ.T.Ü., Design: M.U., İ.T.Ü., Data Collection or Processing: M.K., Analysis or Interpretation: M.K., M.U., İ.T.Ü., Literature Search: M.K., Writing: M.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74:229-63.
2. Deng M, Ma X, Liang X, Zhu C, Wang M. Are pretreatment neutrophil-lymphocyte ratio and platelet-lymphocyte ratio useful in predicting the outcomes of patients with small-cell lung cancer? Oncotarget. 2017;8:37200-7.
3. Nogueiro J, Santos-Sousa H, Pereira A, et al. The impact of the prognostic nutritional index (PNI) in gastric cancer. Langenbecks Arch Surg. 2022;407:2703-14.
4. Gulmez A, Coskun H, Koseci T, Ata S, Bozkurt B, Cil T. Effect of microsatellite status and pan-immune-inflammation score on pathological response in patients with clinical stage III stomach cancer treated with perioperative chemotherapy. Medicina (Kaunas). 2023;59:1625.
5. Yazici H, Yegen SC. Is Systemic inflammatory response index (SIRI) a reliable tool for prognosis of gastric cancer patients without neoadjuvant therapy? Cureus. 2023 Mar 23;15(3):e36597.
6. Hirahara N, Matsubara T, Fujii Y, et al. Comparison of the prognostic value of immunoinflammation-based biomarkers in patients with gastric cancer. Oncotarget. 2020;11:2625-35.
7. Szczepanik AM, Paszko A, Szura M, Scully-Horner T, Kulig J. Alternative staging of regional lymph nodes in gastric cancer. Prz Gastroenterol. 2016;11:145-9.
8. Posteraro B, Persiani R, Dall'Armi V, et al. Prognostic factors and outcomes in Italian patients undergoing curative gastric cancer surgery. Eur J Surg Oncol. 2014;40:345-51.
9. Jiang Q, Zeng X, Zhang C, et al. Lymph node ratio is a prospective prognostic indicator for locally advanced gastric cancer patients after neoadjuvant chemotherapy. World J Surg Oncol. 2022;20:261.
10. Zhou P, Sun X, Zeng L, et al. Lymph node ratio is a prognostic indicator for locally advanced gastric cancer after neoadjuvant immunochemotherapy. BMC Gastroenterol. 2024;24:371.
11. Agalar C, Aysal A, Unek T, et al. The role of log odds of positive lymph nodes in predicting the survival after resection for ampullary adenocarcinoma. Pathol Oncol Res. 2020;26:467-73.
12. Gu P, Deng J, Sun Z, et al. Superiority of log odds of positive lymph nodes (LODDS) for prognostic prediction after gastric cancer surgery: a multi-institutional analysis of 7620 patients in China. Surg Today. 2021;51:101-10.
13. Sun Z, Xu Y, Li de M, et al. Log odds of positive lymph nodes: a novel prognostic indicator superior to the number-based and the ratio-based N category for gastric cancer patients with R0 resection. Cancer. 2010;116:2571-80.
14. Qiu MZ, Qiu HJ, Wang ZQ, et al. The tumor-log odds of positive lymph nodes-metastasis staging system, a promising new staging system for gastric cancer after D2 resection in China. PLoS One. 2012;7:e31736.
15. Aurello P, Petrucciani N, Nigri GR, et al. Log odds of positive lymph nodes (LODDS): what are their role in the prognostic assessment of gastric adenocarcinoma? J Gastrointest Surg. 2014;18:1254-60.
16. He X, Wu W, Lin Z, Ding Y, Si J, Sun LM. Validation of the American Joint Committee on Cancer (AJCC) 8th edition stage system for gastric cancer patients: a population-based analysis. Gastric Cancer. 2018;21:391-400.
17. Ren JY, Xu M, Niu XD, et al. Systemic inflammatory response index is a predictor of prognosis in gastric cancer patients: retrospective cohort and meta-analysis. World J Gastrointest Surg. 2024;16:382-95.
18. Liu Z, Ge H, Miao Z, Shao S, Shi H, Dong C. Dynamic changes in the systemic inflammation response ındex predict the outcome of resectable gastric cancer patients. Front Oncol. 2021;11:577043.
19. Chen J, Hong D, Zhai Y, Shen P. Meta-analysis of associations between neutrophil-to-lymphocyte ratio and prognosis of gastric cancer. World J Surg Oncol. 2015;13:122.
20. Wu J, Wu XD, Gao Y, Gao Y. Correlation between preoperative systemic immune-inflammatory indexes and the prognosis of gastric cancer patients. Eur Rev Med Pharmacol Sci. 2023;27:5706-20.
21. Wang K, Diao F, Ye Z, et al. Prognostic value of systemic immune-inflammation index in patients with gastric cancer. Chin J Cancer. 2017;36:75.
22. Migita K, Takayama T, Saeki K, et al. The prognostic nutritional index predicts long-term outcomes of gastric cancer patients independent of tumor stage. Ann Surg Oncol. 2013;20:2647-54.
23. Xu Z, Chen X, Yuan J, Wang C, An J, Ma X. Correlations of preoperative systematic immuno-inflammatory index and prognostic nutrition index with a prognosis of patients after radical gastric cancer surgery. Surgery. 2022;172:150-9.
24. Attaallah W, Uprak K, Gunal O, Yegen C. Prognostic impact of the metastatic lymph node ratio on survival in gastric cancer. Indian J Surg Oncol. 2016;7:67-72.
25. Ergenç M, Uprak TK, Akın Mİ, Hekimoğlu EE, Çelikel ÇA, Yeğen C. Prognostic significance of metastatic lymph node ratio in gastric cancer: a Western-center analysis. BMC Surg. 2023;23:220.
26. Díaz Del Arco C, Estrada Muñoz L, Sánchez Pernaute A, et al. Prognostic role of the log odds of positive lymph nodes in Western patients with resected gastric cancer: a comparison with the 8th edition of the TNM staging system. Am J Clin Pathol. 2024;161:186-96.

DOI: 10.4274/forbes.galenos.2026.36744
Forbes J Med 2026;7:7-14

Original Article

The Effect of Preoperative Kellgren–Lawrence Grade on Length of Stay and Early Postoperative Complications After Primary Total Knee Arthroplasty

Yılmaz ÖNDER[1], Tolgacan KURTULUŞ[2], Mahmut TUNÇEZ[1], Tuğrul BULUT[1]

[1]İzmir Katip Çelebi University, Atatürk Training and Research Hospital, Department of Orthopaedics and Traumatology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Orthopaedics and Traumatology, İzmir, Türkiye

Cite as: Önder Y, Kurtuluş T, Tunçez M, Bulut T. The effect of preoperative Kellgren–Lawrence grade on length of stay and early postoperative complications after primary total knee arthroplasty. Forbes J Med. 2026;7:7-14

ABSTRACT

Introduction: The aim of this study is to investigate the relationship between preoperative Kellgren–Lawrence (KL) grade and hospital length of stay and early postoperative complications in patients undergoing primary total knee arthroplasty (TKA).

Methods: Patients who underwent cemented posterior-stabilized TKA between January 2023 and January 2026 were retrospectively evaluated. Patients aged ≥55 years with a diagnosis of primary knee osteoarthritis (OA), preoperative weight-bearing knee plain radiographs, and at least 6 months of follow-up were included in the study. Valgus deformity, KL grade 2, body mass index (BMI) ≥35, inflammatory arthritis, post-traumatic arthritis, revision surgery, and patients without adequate follow-up or suitable plain radiographs were excluded. Patients were grouped according to their KL grades. All demographic variables, in addition to hospital length of stay and early minor and major complications, were statistically evaluated. Statistical significance was defined as a p value <0.05 for all analyses.

Results: A total of 482 patients were included in the study (KL grade 3: n=254, 52.7%; KL grade 4: n=228, 47.3%). There were no significant differences between the groups in terms of age, BMI, surgical time, American Society of Anesthesiologists score, Charlson comorbidity index, and blood transfusion requirement (for each p>0.05). Hospital length of stay was significantly longer in the KL grade 4 group (p=0.004). The rate of major complications was higher in the KL grade 4 group (p=0.031), while the rates of minor complications were similar (p>0.05).

Conclusion: Increased preoperative radiographic OA severity is associated with longer hospital length of stay and a higher rate of major complications after primary TKA.

Keywords: Kellgren–Lawrence grade, knee osteoarthritis, total knee arthroplasty, length of stay, early postoperative complication

Received: 05.02.2026
Accepted: 03.03.2026
Epub: 04.03.2026
Publication Date: 10.03.2026

Corresponding Author:
Tuğrul BULUT
İzmir Katip Çelebi University, Atatürk Training and Research Hospital, Department of Orthopaedics and Traumatology, İzmir, Türkiye
drtugrulbulut@yahoo.com
ORCID ID: 0000-0002-7075-0873

INTRODUCTION

In knee osteoarthritis (OA), patients' tendency to avoid surgery and their efforts to delay it by resorting to non-surgical alternative treatments can exacerbate limitations in physical activity, potentially increasing the risk of further disability and chronic disease.[1,2] Consequently, total knee arthroplasty (TKA) is often performed at more advanced radiographic stages, particularly in Kellgren–Lawrence (KL) grade 4. When the relevant literature is reviewed, studies comparing patients with advanced-stage KL grades 3 and 4 report that the two groups are often similar with respect to postoperative functional outcomes and patient-reported scores after TKA.[3,4] However, studies evaluating critical outcomes, such as hospital length of stay and early complications, comparing these two grades have been relatively limited.

In KL grade 3 patients, joint deformities and soft tissue contractures are more limited than in KL grade 4 patients. This can affect the level of surgical difficulty, the early postoperative recovery process, and the risk of complications. Therefore, comparing early hospital outcomes in patients with KL grades 3 and 4 is important

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

clinically and for determining their burden on the healthcare system.

The hypothesis of this study is that patients with KL grade 3 will have a shorter hospital length of stay and lower rates of early postoperative complications compared to patients with KL grade 4.

The aim of this study is to investigate the relationship between preoperative KL grade, hospital length of stay, and early postoperative complication rates in patients undergoing primary TKA.

METHODS

This retrospective study was approved by the İzmir Katip Çelebi University Institutional Review Board (approval number: 0015, date: 15.01.2026). All procedures involving human participants were conducted in accordance with the institutional committee's ethical standards.

Patient Inclusion

In this study, 1057 patients who underwent primary TKA between January 2023 and January 2026 were retrospectively evaluated. Patients who underwent cemented posterior-stabilized TKA for KL grade 3 or 4 primary knee OA; who were ≥55 years of age, had preoperative weight-bearing knee plain radiographs, and had at least 6 months of postoperative follow-up been included in the study. Exclusion criterias for this study were; revision TKA, valgus deformity OA, KL grade 2 OA, body mass index (BMI) ≥35, uncemented or cruciate-retaining TKA, inflammatory arthritis, post-traumatic osteoarthritis, concurrent major joint surgery, tumor or active infection, and lack of suitable radiographs or adequate postoperative follow-up. In total, 482 patients were included in the study (Figure 1).

Surgical Technique and Follow-Up Protocol

All patients underwent surgery under regional anesthesia. A midline skin incision was made with a pneumatic tourniquet, and a medial arthrotomy was performed via a standard medial parapatellar approach. Cemented posterior-stabilized TKA was performed without patellar resurfacing. The hemovac drain was removed on the first postoperative day. Mobilization and active range-of-motion exercises were initiated with full weight-bearing, using walking aids. 5000 IU of low-molecular-weight heparin was administered subcutaneously for 30 days as venous thromboembolism prophylaxis. As antibiotic prophylaxis, 1 g of cefazolin was administered intravenously as a single dose before surgery, followed by three doses within the first 16 hours after surgery. Patients were discharged when pain was controlled with oral medication, functional independence was achieved with walking aids, and prolonged wound drainage requiring hospital follow-up was no longer present. Clinical and radiological follow-ups were performed weekly during the first month, monthly for the next 5 months, and quarterly thereafter.

Patient Evaluation

Detailed demographic and clinical data, such as age, gender, BMI, American Society of Anesthesiologists (ASA) score, Charlson comorbidity index (CCI), surgical time, blood transfusion requirement, and hospital length of stay, were recorded for all patients. Minor and major complications developing during follow-up periods of at least 6 months were recorded. Complications were classified as minor or major according to their clinical severity and the level of treatment required. Classification was performed in accordance with widely accepted surgical complication grading systems and the TKA literature.[5-8]

Minor complications were defined as events that could be managed with conservative treatment or medical support without requiring invasive surgery or intensive care. This group included prolonged serous wound drainage (discharge lasting longer than 72 hours and causing more than 2x2 cm of wetting in the dressing), superficial wound infection (limited to the skin and subcutaneous tissue and resolving with antibiotic treatment), hematoma or seroma (not requiring surgical drainage), limitation of range of motion (ROM), urinary retention or urinary tract infection, electrolyte imbalances, and transient gastrointestinal complications.

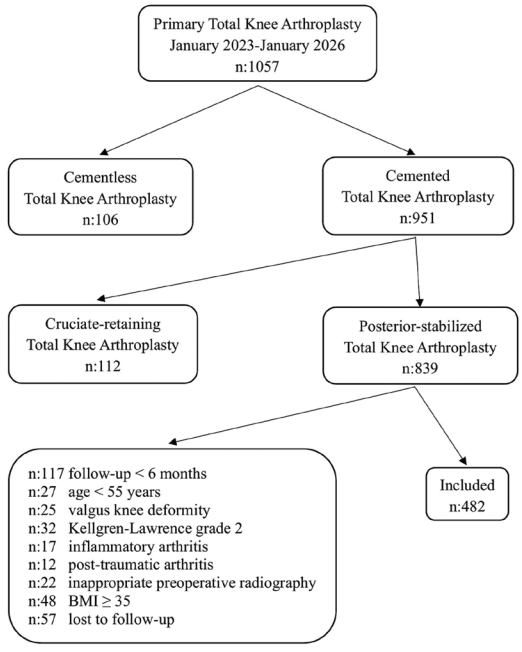


Figure 1. Patient flowchart of the study

BMI: Body mass index

Major complications were defined as life-threatening events requiring surgical, endoscopic, or interventional treatment; requiring intensive care monitoring; or associated with permanent morbidity or mortality. This group included the following: periprosthetic joint infection (requiring surgical debridement, component exchange, or revision); superficial wound necrosis (requiring debridement); intraoperative vascular, nerve, or ligament injuries; venous thromboembolism (deep vein thrombosis or pulmonary embolism); major cardiovascular diseases (myocardial infarction, stroke); acute renal failure; sepsis or septic shock; periprosthetic fractures; complications requiring reoperation such as dislocation; and postoperative mortality.

Radiographic Evaluation

Radiological assessments were performed digitally using the picture archiving and communication system. All patients were evaluated according to the KL grading system using preoperative weight-bearing anteroposterior and lateral plain radiographs of the knee.[9] The presence of numerous osteophytes, significant joint space narrowing, sclerosis, and possible bone deformity was classified as KL grade 3 (Figure 2). Large osteophytes, significant joint space narrowing, severe sclerosis, and definite bone deformity were classified as KL grade 4 (Figure 3). All evaluations were conducted by two independent observers — one orthopedic specialist and one radiologist — and decisions were reached by consensus.

Statistical Analysis

Data were analyzed using the IBM SPSS Statistics Standard Concurrent User V 22 (IBM Corp, Armonk, New York, USA) statistical software package. Descriptive statistics were presented as number of units (n), percentage (%), mean ± standard deviation ($\bar{x}$ ± SD), median, minimum, and maximum (max) values. Normality of the numerical variables was assessed using the Kolmogorov-Smirnov test. The numerical variables did not conform to the assumption of normality. The Mann-Whitney U test was used to compare numerical variables among KL grade groups. The Pearson chi-square test was used to compare categorical variables across KL grades. The statistical significance level was set at $p<0.05$.

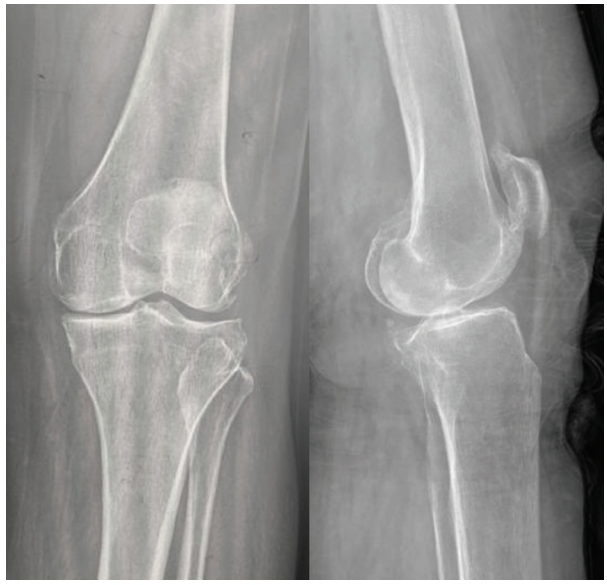


Figure 2. Weight-bearing anteroposterior and lateral knee plain radiographs of a 62-year-old female patient with Kellgren–Lawrence grade 3 knee osteoarthritis



Figure 3. Weight-bearing anteroposterior and lateral knee radiographs of Kellgren–Lawrence grade 4 knee osteoarthritis: (a) 65-year-old female patient; (b) 68-year-old female patient. Significant joint deformity and flexion contracture are observed in (b)

RESULTS

The study included 254 patients with KL grade 3 (52.7%) and 228 patients with KL grade 4 (47.3%). Three hundred and ninety-eight (82.6%) of the patients were female and 84 (17.4%) were male. When all patients were evaluated, the mean age was 67.7±6.7 years, the mean follow-up was 19.4±8.7 months, the mean BMI was 26.6±2.0 kg/m², the mean surgical time was 100.5±24.6 minutes, and the mean hospital length of stay was 5.8±2.2 days. The mean transfusion amount in patients who received blood transfusions was 1.9±0.9 units. When complications were evaluated, no systemic complications affecting major organ systems were detected in any patient. Minor complications were observed in 48 patients (10.0%), and major complications in 54 patients (11.2%) (Table 1).

Prolonged wound drainage developed in 18 patients (3.7%). These cases were managed without surgery by temporarily discontinuing prophylactic low-molecular-weight heparin and repeating dressing changes. Superficial wound infection occurred in 20 patients (4.2%) and was successfully treated with oral antibiotic therapy and local wound care. Eight hematomas (1.7%) that developed in the postoperative period were managed conservatively by aspiration. Two patients (0.4%) experienced a postoperative restricted ROM. One of these patients showed clinical improvement with physical therapy, whereas the other did not. In this case, a 10° extension restriction developed during follow-up, and max knee flexion was limited to 90° (Table 2).

Among major complications, deep vein thrombosis developed in 1 patient (0.2%) and was successfully managed with medical treatment. Intraoperative medial collateral ligament injury was detected in 4 patients (0.8%); these cases were successfully treated with primary repair using screw suture anchors and knee brace immobilization for 6 weeks. In 2 patients (0.4%), surgical debridement was performed because of superficial wound necrosis, and negative-pressure wound therapy, also known as vacuum-assisted closure, was used to promote wound healing. One patient (0.2%) experienced a dislocation following a fall down the stairs during the postoperative period and underwent revision surgery. Additionally, one patient (0.2%) developed an intraoperative tibial periprosthetic fracture; internal fixation was achieved in the same surgical session. Regarding infectious complications, 29 cases (6.0%) that developed early acute infection were treated with debridement, antibiotics, and implant retention with exchange of the polyethylene insert. One patient (0.2%) underwent one-stage revision arthroplasty due to an infection that developed during the second postoperative month. Two-stage revision arthroplasty was performed in 15 patients (3.2%) (Table 2).

No statistically significant differences were found between the groups included in the study with respect to age, gender, follow-up, side (right/left), blood transfusion, surgical time, BMI, ASA score, CCI, and presence of minor complications (for each $p>0.05$) (Table 3).

The presence of major complications differed significantly across KL grades ($p=0.031$). It was found that patients in KL grade 4 developed major complications more frequently than those in KL grade 3. In addition, hospital length of stay showed a statistically significant difference between KL grades ($p=0.004$). It was found that the hospital length of stay was longer in KL grade 4 patients compared to KL grade 3 patients (6.1±2.7 days; 5.5±1.6 days, respectively) (Table 3).

Table 1. Descriptive variables of patients

Variables	n	%
Sex		
Female	398	82.6
Male	84	17.4
Side		
Right	233	48.3
Left	249	51.7
ASA score		
1	11	2.2
2	374	77.5
3	97	20.3
CCI		
0-2	57	11.8
3-4	280	58.1
≥5	145	30.1
KL grade		
3	254	52.7
4	228	47.3

Table 1. Continued		
Variables	**n**	**%**
Blood transfusion		
Negative	180	37.3
Positive	302	62.7
Blood transfusion (unit)		
0	180	37.3
1	124	25.7
2	123	25.5
3	38	7.9
4	10	2.1
5	7	1.5
Minor complication		
Negative	434	90.0
Positive	48	10.0
Major complication		
Negative	428	88.8
Positive	54	11.2
	$\bar{x}$ ± SD	**M (min-max)**
Age (year)	67.7±6.7	68.0 (55.0-87.0)
Follow-up (month)	19.4±8.7	18.4 (6.3-36.2)
BMI	26.6±2.0	27.0 (21.0-34.0)
Surgical time (minute)	100.5±24.6	100.0 (40.0-190.0)
Length of stay (day)	5.8±2.2	5.0 (3.0-27.0)
Blood transfusion (unit)	1.9±0.9	2.0 (1.0-5.0)

n: Number of patients, %: Column percentage, $\bar{x}$ ± SD: Mean ± standard deviation, M: Median, min: Minimum, max: Maximum, BMI: Body mass index, ASA: American Society of Anesthesiologists, CCI: Charlson comorbidity index, KL: Kellgren–Lawrence

Table 2. Distribution of complications in groups						
	All patients		**KL grade 3**		**KL grade 4**	
	n	**%**	**n**	**%**	**n**	**%**
Minor complication	48	10.0	23	9.0	25	11.0
Prolonged wound drainage	18	3.7	11	4.3	7	3.1
Superficial wound infection	20	4.2	8	3.1	12	5.3
Hematoma	8	1.7	3	1.2	5	2.2
Restricted ROM	2	0.4	1	0.4	1	0.4
Major complication	54	11.2	21	8.3	33	14.5
Superficial wound necrosis	2	0.4	-	-	2	0.9
Infection (DAIR)	29	6.0	12	4.7	17	7.5
Infection (two-stage revision)	15	3.2	5	2.0	10	4.4
Infection (one-stage revision)	1	0.2	1	0.4	-	-
DVT	1	0.2	-	-	1	0.4
Dislocation	1	0.2	-	-	1	0.4
Intraoperative periprosthetic fracture	1	0.2	1	0.4	-	-
Intraoperative MCL injury	4	0.8	2	0.8	2	0.9

KL: Kellgren–Lawrence, ROM: Range of motion, DAIR: Debridement, antibiotic, and implant retention, DVT: Deep vein thrombosis, MCL: Medial collateral ligament

Table 3. Comparison of variables according to groups					
Variables	**KL grade 3**		**KL grade 4**		**p value**
	n	**%**	**n**	**%**	
Sex					
Sex	213	83.9	185	81.1	0.432
Male	41	16.1	43	18.9	
Side					
Right	129	50.8	104	45.6	0.256
Left	125	49.2	124	54.4	
ASA score					
1	8	3.1	3	1.3	0.336
2	170	66.9	176	77.2	
3	76	29.9	49	21.5	
CCI					
0-2	24	9.4	33	14.5	0.152
3-4	147	57.9	133	58.3	
≥5	83	32.7	62	27.2	
Blood transfusion (unit)					
0	96	37.8	84	36.8	0.646
1	68	26.8	56	24.6	
2	59	23.2	64	28.1	
≥3	31	12.2	24	10.5	
Blood transfusion					
Negative	96	37.8	84	36.8	0.829
Positive	158	62.2	144	63.2	
Minor complication					
Negative	231	90.9	203	89.0	0.485
Positive	23	9.1	25	11.0	
Major complication					
Negative	233	91.7	195	85.5	0.031
Positive	21	8.3	33	14.5	
	$\bar{x}$ ± SD	**M (min-max)**	**$\bar{x}$ ± SD**	**M (min-max)**	
Age (year)	67.8±6.6	68.0 (56.0-85.0)	67.7±6.9	68.0 (55.0-87.0)	0.830
Follow-up (month)	19.0±9.4	16.9 (6.3-36.2)	19.9±7.8	19.5 (6.5-35.9)	0.105
BMI	26.7±2.0	27.0 (21.0-34.0)	26.5±2.0	27.0 (21.0-29.0)	0.592
Surgical time (minute)	99.7±26.8	100.0 (40.0-190.0)	101.3±22.0	105.0 (45.0-160.0)	0.202
Length of stay (day)	5.5±1.6	5.0 (3.0-15.0)	6.1±2.7	5.0 (3.0-27.0)	0.004
Blood transfusion (unit)	1.9±1.0	2.0 (1.0-5.0)	1.8±0.8	2.0 (1.0-5.0)	0.820

KL: Kellgren–Lawrence, n: Number of patients, %: Column percentage, ASA: American Society of Anesthesiologists, CCI: Charlson comorbidity index. $\bar{x}$ ± SD: Mean ± standard deviation, M: median, min: Minimum, max: Maximum, BMI: Body mass index

DISCUSSION

The results demonstrated that the preoperative radiographic severity of OA in patients undergoing primary TKA was significantly associated with length of hospital stay and early postoperative complications. Hospital length of stay was longer in KL grade 4 patients, and the rate of early-period major complications was significantly higher than that in KL grade 3 patients. One possible explanation for this relationship is that the KL classification exhibits a pronounced ceiling effect in advanced knee OA and does not provide sufficient detail for severe cases. KL grade 4 defines advanced disease, in which the joint space is completely lost; however, it does not sufficiently differentiate among parameters such as the degree of deformity, bone loss, osteophyte burden, and soft-tissue contracture, which can directly affect surgical difficulty and early rehabilitation. Therefore, it is possible for a patient with a completely closed joint space but minimal deformity to be assigned the same grade (KL 4) as a patient with advanced deformity and severe contracture. Keenan et al.[10] have highlighted the limited discriminatory power of the KL grading system in advanced OA. Although the degrees of deformity and soft tissue contracture were not objectively measured in our study, it is likely that cases with more complex deformities and contractures were relatively more prevalent in the KL grade 4 group. In this patient group, more challenging surgical exposure, increased need for ligament balancing, and delayed postoperative rehabilitation may be associated with prolonged hospital stay and increased risk of early postoperative complications.

The observation that KL grade 4 patients had longer hospital stays in our study is an important finding with significant implications for healthcare resource utilisation and costs. A review of the relevant literature shows that prolonged hospital stay is associated with increased rates of complications and readmissions following joint arthroplasty.[10-13] The most important reasons for this are the frequent occurrence of marked joint deformities, widespread osteophyte formation, subchondral bone changes, and capsular contractures in KL grade 4 knee OA. These structural changes can make the surgical procedure more complex from a technical point of view. As a result, postoperative pain control may become more difficult, and early mobilisation may be delayed. The time to achieve functional independence may be prolonged. Consequently, an increase in hospital length of stay may be anticipated. None of the patients in our study experienced complications that prolonged length of hospital stay, except for prolonged wound drainage. All complications, except for medial collateral ligament injury and periprosthetic fracture, developed after discharge. Considering that prolonged wound drainage is a minor complication treated conservatively, the degree of OA is responsible for the longer hospital stay.

Another important finding was that the rate of early major complications was significantly higher among patients with KL grade 4. We believe that there are two reasons for this. The expected progression of OA to the final radiographic grade may increase joint deformity and capsular contracture, thereby making surgery more difficult. Second, in advanced OA, the decline in joint-related functional capacity and lower-extremity muscle strength may lead to deterioration in patients' overall health, making them more susceptible to complications. Indeed, studies have shown that advanced knee OA is associated with frailty syndrome, characterised by a decline in physical reserves and reduced physiological resilience to stress, and may increase frailty.[14] Although traditional risk classification systems such as ASA and CCI reflect the general health status of patients, an advanced radiographic stage of OA may increase the risk of early complications independent of the burden of systemic comorbidity. It should be borne in mind that these indices, which are frequently used in clinical practice, may not be sufficient on their own to predict the risk of early postoperative complications after TKA.

In recent years, insurance companies in many developed countries, particularly the United States, have introduced pre-authorisation criteria for TKA and restricted surgery to patients with KL grade 4 OA.[4] This approach is based on the assumption that the clinical benefit of surgery may be lower for patients with early-grade OA.[15,16] However, our study shows that delaying surgery until the radiographic end-stage may create disadvantages in terms of hospital length of stay and early major complications. The KL classification system cannot clearly distinguish the severity within advanced OA. Therefore, we believe that radiographic grade should not be the sole criterion for surgical decision-making. This approach may lead to the postponement of definitive treatment, particularly in symptomatic KL grade 3 patients. This is an important parameter that may affect early hospital-based outcomes. Our study shows that preoperative radiographic OA severity is associated with early hospital-based outcomes after primary TKA. Longer hospital stays and higher rates of early-period major complications have been observed among KL grade 4 patients. These findings suggest that an advanced radiographic stage may be a clinical indicator affecting not only disease severity but also the early postoperative recovery process and hospital resource utilisation.

This study has some limitations. The retrospective design precludes establishing causal relationships, and unmeasured confounding variables may have influenced the results. The lack of patient-reported outcome measures and long-term functional outcome assessments limits the study to early hospital-based outcomes. Furthermore, the lack of assessment of the inter- and intra-observer reliability of radiographic evaluations may increase the risk of observer bias. Nevertheless, the analysis of a large patient cohort, a homogeneous surgical procedure, and standardised early-period outcomes is among the strengths of this study.

CONCLUSION

This study demonstrates that the preoperative KL grade in patients undergoing primary TKA may create clinically significant differences in hospital length of stay and early postoperative major complications. Patients with KL grade 4 OA were found to have longer hospital stays and higher rates of early major complications. These findings suggest that severe radiographic OA may increase surgical complexity and impair early postoperative recovery, independent of systemic comorbidity scores. Prospective, multicentre studies will support the evaluation of advanced-stage OA through the use of more detailed radiographic classifications and the development of clinical decision algorithms for optimal timing of surgery.

Ethics

Ethics Committee Approval: This retrospective study was approved by the İzmir Katip Çelebi University Institutional Review Board (approval number: 0015, date: 15.01.2026).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Y.Ö., M.T., T.K., T.B., Concept: Y.Ö., M.T., T.B., Design: Y.Ö., M.T., T.K., Data Collection or Processing: Y.Ö., T.K., Analysis or Interpretation: Y.Ö., T.B., Literature Search: Y.Ö., T.B., Writing: Y.Ö., M.T., T.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Ravi B, Croxford R, Austin PC, et al. The relation between total joint arthroplasty and risk for serious cardiovascular events in patients with moderate-severe osteoarthritis: propensity score matched landmark analysis. BMJ. 2013;347:f6187.
2. Turkiewicz A, Kiadaliri AA, Englund M. Cause-specific mortality in osteoarthritis of peripheral joints. Osteoarthritis Cartilage. 2019;27:848-54.
3. Scott CEH, Holland G, Keenan OJF, et al. Radiographic severity, extent and pattern of cartilage loss are not associated with patient reported outcomes before or after total knee arthroplasty in end-stage knee osteoarthritis. Knee. 2021;31:54-63.
4. Goh GS, Schwartz AM, Friend JK, et al. Patients who have Kellgren-Lawrence grade 3 and 4 osteoarthritis benefit equally from total knee arthroplasty. J Arthroplasty. 2023;38:1714-7.
5. McNally M, Sousa R, Wouthuyzen-Bakker M, et al. The EBJIS definition of periprosthetic joint infection. Bone Joint J. 2021;103:18-25.
6. Ghanem E, Heppert V, Spangehl M, et al. Wound management. J Orthop Res. 2014;32:S108-19.
7. Simons MJ, Amin NH, Scuderi GR. Acute wound complications after total knee arthroplasty: prevention and management. J Am Acad Orthop Surg. 2017;25:547-55.
8. Warren J, Sundaram K, Anis H, et al. Spinal anesthesia is associated with decreased complications after total knee and hip arthroplasty. J Am Acad Orthop Surg. 2020;28:e213-e21.
9. Kohn MD, Sassoon AA, Fernando ND. Classifications in brief: Kellgren-Lawrence classification of osteoarthritis. Clin Orthop Relat Res. 2016;474:1886-93.
10. Keenan OJF, Holland G, Maempel JF, Keating JF, Scott CEH. Correlations between radiological classification systems and confirmed cartilage loss in severe knee osteoarthritis. Bone Joint J. 2020;102:301-9.
11. Otero JE, Gholson JJ, Pugely AJ, Gao Y, Bedard NA, Callaghan JJ. Length of hospitalization after joint arthroplasty: does early discharge affect complications and readmission rates? J Arthroplasty. 2016;31:2714-25.
12. DeMik DE, Carender CN, An Q, Callaghan JJ, Brown TS, Bedard NA. Longer length of stay is associated with more early complications after total knee arthroplasty. Iowa Orthop J. 2022;42:53-9.
13. Sarpong NO, Boddapati V, Herndon CL, Shah RP, Cooper HJ, Geller JA. Trends in length of stay and 30-day complications after total knee arthroplasty: an analysis from 2006 to 2016. J Arthroplasty. 2019;34:1575-80.
14. Joo SH, Song JW, Shin K, Kim MJ, Lee J, Song YW. Knee osteoarthritis with a high grade of Kellgren-Lawrence score is associated with a worse frailty status, KNHANES 2010-2013. Sci Rep. 2023;13:19714.
15. Polkowski GG 2nd, Ruh EL, Barrack TN, Nunley RM, Barrack RL. Is pain and dissatisfaction after TKA related to early-grade preoperative osteoarthritis? Clin Orthop Relat Res. 2013;471:162-8.
16. Leppänen S, Niemeläinen M, Huhtala H, Eskelinen A. Mild knee osteoarthritis predicts dissatisfaction after total knee arthroplasty: a prospective study of 186 patients aged 65 years or less with 2-year follow-up. BMC Musculoskelet Disord. 2021;22:657.

DOI: 10.4274/forbes.galenos.2026.87609
Forbes J Med 2026;7:15-22

Original Article

Biochemical and Hematological Markers for Predicting Difficult Laparoscopic Cholecystectomy in Patients Aged ≥65 Years: A Retrospective Cohort Study

Ozan Barış NAMDAROĞLU[1], Ferhat DEMİRCİ[2], Fatma DİKİŞER[1], Beyza GÜMÜŞTEKİN[1]

[1]University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Clinic of General Surgery, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Clinic of Medical Biochemistry, İzmir, Türkiye

Cite as: Namdaroğlu OB, Demirci F, Dikişer F, Gümüştekin B. Biochemical and hematological markers for predicting difficult laparoscopic cholecystectomy in patients aged ≥65 years: a retrospective cohort study. Forbes J Med. 2026;7:15-22

ABSTRACT

Objective: To identify factors associated with difficult laparoscopic cholecystectomy (DLC) in patients aged 65 years and older and to evaluate the predictive value of inflammatory markers.

Methods: This single-center retrospective cohort study included patients aged 65 years and older who underwent laparoscopic cholecystectomy between 2015 and 2025. Difficult surgery was defined as conversion to open surgery and/or subtotal (bail-out) cholecystectomy and/or operative time of 120 minutes or longer. The C-reactive protein-to-albumin (CAR) ratio was calculated as C-reactive protein (CRP) divided by albumin. Multivariable logistic regression was used to identify independent predictors.

Results: A total of 726 patients were analyzed; the median age was 70 years (25^{th}–75^{th} percentile, 67–74), and 35.7% were male. Surgical difficulty occurred in 276 patients (38.0%). Conversion to open surgery occurred in 95 patients (13.1%) and was most commonly due to unsafe or uncertain anatomy (89.5% of conversions). The difficult group had a longer hospital stay (median 4 days compared with 1 day), a greater need for intensive care (34.8% compared with 8.4%), and a higher 30-day mortality (2.9% compared with 0.4%). Independent predictors were: acute cholecystitis [adjusted odds ratio (aOR) 8.79; 95% confidence interval (CI): 2.00–38.73]; a higher log-transformed CAR ratio (aOR): 1.45; 95% CI: 1.29–1.63); and male sex aOR: 1.44; 95% CI: 1.02–2.03).

Conclusion: In older adults, DLC are common and associated with worse perioperative outcomes. The CAR ratio may support preoperative risk stratification and operative planning in this population.

Keywords: Cholecystectomy, aged, C-reactive protein, albumin

Received: 16.02.2026
Accepted: 04.03.2026
Epub: 06.03.2026
Publication Date: 13.03.2026

Corresponding Author:
Ozan Barış NAMDAROĞLU
University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Department of General Surgery, İzmir, Türkiye
ozannamdaroglu@yahoo.com
ORCID ID: 0000-0001-5195-307X

INTRODUCTION

Laparoscopic cholecystectomy is the standard surgical approach for symptomatic gallstone disease. In older patients, a higher burden of comorbidities and a potentially more severe inflammatory course may increase technical difficulty, including adhesions, edema, and obscured anatomy. These factors can prolong operative time and increase the need for bail-out procedures or conversion to open surgery.[1-3]

The consequences of a difficult cholecystectomy extend beyond technical complexity and may translate into clinically meaningful outcomes, such as increased intraoperative complications, postoperative morbidity, intensive care unit (ICU) requirements, and longer hospital stays.[1,2] Therefore, objective preoperative risk estimation—particularly in patients aged ≥65 years—is important for operating room planning, anticipating bail-out strategies, and improving preoperative counseling.

Several clinical scoring systems and imaging-based predictors have been proposed to identify difficult cholecystectomies. However, in real-world retrospective datasets, these variables are not always recorded in a standardized or sufficiently granular manner, limiting

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

their consistent use, especially in large cohorts. For this reason, we adopted an approach based on objective and reproducible laboratory parameters routinely obtained at initial presentation.

C-reactive protein (CRP) reflects the acute-phase response, whereas albumin is inversely related to inflammation and serves as a marker of overall physiological reserve. The CRP/to-albumin ratio (CAR), calculated as CRP (mg/L) divided by albumin (g/dL), combines both parameters into a single measure and has been evaluated as a predictor of difficult laparoscopic cholecystectomy (DLC) in patients with acute cholecystitis.[4] The neutrophil-to-lymphocyte ratio (NLR) is another practical marker of systemic inflammatory response and has been associated with cholecystitis severity.[5] The aim of this study was to identify clinical features associated with difficult surgery during laparoscopic cholecystectomy in patients aged ≥65 years and to assess the predictive performance of readily available laboratory markers such as CAR and NLR.

METHODS

Study Design and Patient Selection

This was a single-center, retrospective, observational cohort study. Consecutive patients aged ≥65 years who underwent laparoscopic cholecystectomy at University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital between January 2015 and June 2025 were evaluated. Data were extracted from electronic medical records and operative/anesthesia notes.

Inclusion criteria were age ≥65 years and patients in whom laparoscopic cholecystectomy was initiated as the primary surgical procedure during the study period, regardless of intraoperative conversion to open surgery or subtotal (bail-out) cholecystectomy.

Exclusion criteria included primary open cholecystectomy (n=16), laparoscopic cholecystectomy performed concomitantly with another major abdominal procedure, missing CRP or albumin values at admission (n=8), and incomplete operative records (n=3) (Figure 1).

Demographic characteristics (age, sex), urgency (elective vs. emergency), comorbidities, American Society of Anesthesiologists (ASA) class, anticoagulant use, laboratory parameters (complete blood count and biochemistry), and perioperative clinical/surgical variables were recorded. Laboratory values were obtained from blood samples collected at the time of initial presentation. Anticoagulant use included vitamin K antagonists (warfarin), direct oral anticoagulants (apixaban, rivaroxaban, dabigatran), and antiplatelet agents (acetylsalicylic acid and P2Y12 inhibitors such as clopidogrel), all managed according to institutional perioperative protocols. CAR was calculated as CRP (mg/L)/albumin(g/dL) and NLR was calculated as neutrophil/lymphocyte ratio. All data were anonymized prior to analysis.

Surgical Approach and Definitions

Indications were categorized as biliary colic, biliary pancreatitis, and acute cholecystitis. The initial surgical approach was laparoscopic.

Difficult cholecystectomy was defined using a composite outcome of any of the following: (i) conversion to open surgery and/or (ii) subtotal (bail-out) cholecystectomy and/or (iii) operative time ≥120 minutes. Operative time has frequently been used as an objective surrogate marker of technical difficulty during laparoscopic cholecystectomy because there is no universal definition of a "difficult" laparoscopic cholecystectomy. Thresholds ranging from 90 to 120 minutes have been adopted in previous studies.[6,7] As our institution is a tertiary referral teaching hospital with resident involvement in surgical training, we selected ≥120 minutes as a conservative and clinically meaningful cut-off to reflect substantial operative prolongation beyond routine cases while minimizing misclassification of moderately extended procedures.

Reasons for conversion were extracted from free-text operative notes and categorized as unsafe/uncertain anatomy, bleeding/hemostasis, biliary tract injury, bowel injury, or other.

Intraoperative complications were recorded. Postoperative complications were graded using the Clavien–Dindo classification.[8,9] Length of hospital stay, ICU requirement, and 30-day mortality were considered secondary outcomes.

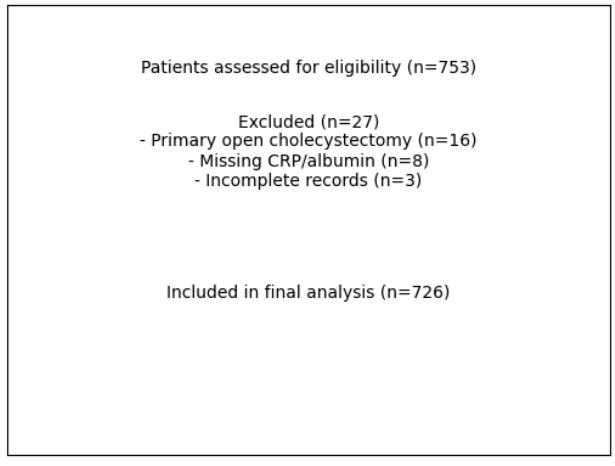


Figure 1. Flow diagram of patient selection and exclusion process

CRP: C-reactive protein

Statistical Analysis

Statistical analyses were performed using SPSS version 28.0 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as median [interquartile range, (IQR)], and categorical variables are presented as n (%). Comparisons between difficult and non-difficult groups were conducted using the Mann–Whitney U test for continuous variables and the chi-square test or Fisher's exact test, as appropriate, for categorical variables.

A multivariable logistic regression model was built to identify independent factors associated with difficult cholecystectomy. Covariates included age, sex, emergency surgery, acute cholecystitis, ASA ≥III, anticoagulant use, log-transformed CAR (logCAR), and log-transformed NLR. Discrimination was assessed using the area under the receiver operating characteristic curve (ROC) area under the curve (AUC). ROC analyses were performed for CAR, CRP, and NLR, and optimal cut-offs were determined using the Youden index. Statistical significance was set at $p<0.05$.

Ethical Approval

Ethics approval was obtained from the University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital Ethics Committee (approval number: 2025/12-28, date: 05.01.2026). The study was conducted in accordance with the Declaration of Helsinki and its later amendments. Informed consent was waived because the study was retrospective.

RESULTS

A total of 726 patients were included (Table 1). The median age was 70 years (IQR 67–74), and 259/726 (35.7%) were male. Seventy procedures (70/726; 9.6%) were performed under emergency conditions. The indications were biliary colic (642/726; 88.4%), acute cholecystitis (66/726; 9.1%), and biliary pancreatitis (18/726; 2.5%).

Difficult cholecystectomy (conversion, subtotal cholecystectomy, or operative time ≥120 minutes) occurred in 276/726 (38.0%) patients. The conversion rate was 95/726 (13.1%), the subtotal (bail-out) rate was 8/726 (1.1%), and operative time ≥120 minutes occurred in 243/726 (33.5%). Conversion was most commonly due to unsafe or uncertain anatomy (85/95; 89.5%; Table 1).

Compared with the non-difficult group, the difficult group had a higher proportion of males [123/276 (44.6%) vs. 136/450 (30.2%); $p<0.001$], emergency surgery [54/276 (19.6%) vs. 16/450 (3.6%); $p<0.001$], acute cholecystitis as the indication [54/276 (19.6%) vs. 12/450 (2.7%); $p<0.001$], ASA ≥III [82/276 (29.7%) vs. 63/450 (14.0%); $p<0.001$], and anticoagulant use [43/276 (15.6%) vs. 30/450 (6.7%); $p<0.001$] (Table 2). Inflammatory markers were also higher in the difficult group, including CRP [12 (4.5–45) vs. 4.5 (1–12); $p<0.001$] and CAR [3.16 (1.12–12.86) vs. 1.07 (0.26–2.86); $p<0.001$].

Operative time was longer in the difficult group [135 (120–160) vs. 90 (74–99) minutes; $p<0.001$]. Intraoperative complications occurred only in the difficult group [14/276 (5.1%) vs. 0/450 (0.0%); $p<0.001$]. The difficult group also had a longer hospital stay [4 (2–7) vs. 1 (1–2) days; $p<0.001$], a higher ICU requirement [96/276 (34.8%) vs. 38/450 (8.4%); $p<0.001$], and a higher 30-day mortality [8/276 (2.9%) vs. 2/450 (0.4%); $p=0.008$].

In multivariable logistic regression (n=726; AUC =0.727), acute cholecystitis [adjusted odds ratio (aOR) 8.79; 95% confidence interval (CI): 2.00–38.73; $p=0.004$], higher logCAR (aOR 1.45; 95% CI 1.29–1.63; $p<0.001$), and male sex (aOR 1.44; 95% CI 1.02–2.03; $p=0.039$) were independently associated with difficult cholecystectomy (Table 3). The Clavien–Dindo distribution differed between groups (overall $p<0.001$; Table 4), and major complications were more frequent in the difficult group [16/276 (5.8%) vs. 5/450 (1.1%); $p<0.001$].

In ROC analysis, CAR showed modest discrimination in predicting difficult cholecystectomy (AUC 0.700), with an optimal cut-off of 2.75 (sensitivity 55.8%, specificity 74.0%). CRP performed similarly (AUC 0.695; cut-off 9.5 mg/L; sensitivity 58.7%; specificity 70.2%), whereas NLR showed limited discrimination (AUC 0.564).

DISCUSSION

There were two main reasons for centering the study on a laboratory-based approach: First, the technical difficulty during cholecystectomy is largely determined by the severity of inflammation and the resulting edema and fibrosis in Calot's triangle, and this biological burden can be objectively represented by CRP, albumin, and their derived ratios. Secondly, routine biochemistry and hemogram parameters are rapid to obtain, inexpensive, and reproducible, even in emergency situations, and partially offset inter-center variability in clinical evaluations or imaging interpretations. We targeted elderly patients (≥65 years) because comorbidities and decreased physiological reserve are more pronounced in this group, clinical findings are sometimes "subdued," and even minor perioperative stressors can more easily lead to functional decline, complications, and mortality. Therefore, rapid and objective stratification of preoperative risk in the elderly population is increasingly clinically important.[10,11]

Since there is no single universal definition of "DLC" in the literature, we used a clinically significant composite endpoint to increase comparability between studies: conversion and/or subtotal (bail-out) cholecystectomy

Table 1. Baseline characteristics, laboratory parameters, operative details, and outcomes (overall cohort)	
Characteristic	**Overall**
Patients, n	726
Age, years, median (IQR)	70 (67–74)
Male sex, n (%)	259/726 (35.7%)
Urgent surgery, n (%)	70/726 (9.6%)
Indication: biliary colic, n (%)	642/726 (88.4%)
Indication: acute cholecystitis, n (%)	66/726 (9.1%)
Indication: biliary pancreatitis, n (%)	18/726 (2.5%)
ASA I, n (%)	122/726 (16.8%)
ASA II, n (%)	459/726 (63.2%)
ASA III, n (%)	130/726 (17.9%)
ASA IV, n (%)	15/726 (2.1%)
ASA V, n (%)	0/726 (0.0%)
Diabetes mellitus, n (%)	177/726 (24.4%)
Hypertension, n (%)	371/726 (51.1%)
COPD, n (%)	52/726 (7.2%)
Coronary artery disease, n (%)	109/726 (15.0%)
Chronic kidney disease, n (%)	11/726 (1.5%)
Anticoagulant use, n (%)	73/726 (10.1%)
WBC (×10^3/μL), median (IQR)	7.7 (6.4–9.3)
CRP (mg/L), median (IQR)	6 (2.2–20)
Albumin (g/dL), median (IQR)	4 (3.6–4.3)
CAR (CRP/albumin), median (IQR)	1.58 (0.55–5)
NLR (neutrophil/lymphocyte), median (IQR)	2.26 (1.73–3.33)
AST (U/L), median (IQR)	22 (19–30)
ALT (U/L), median (IQR)	19 (15–27)
AST/ALT ratio, median (IQR)	1.11 (0.91–1.39)
GGT (U/L), median (IQR)	32 (20–56)
Total bilirubin (mg/dL), median (IQR)	0.7 (0.54–0.9)
Direct bilirubin (mg/dL), median (IQR)	0.11 (0.09–0.18)
Direct/total bilirubin ratio, median (IQR)	0.17 (0.14–0.22)
Operative time (min), median (IQR)	99 (85–125)
Conversion to open, n (%)	95/726 (13.1%)
Conversion reasons among converted cases	(n=95)
Unsafe/unclear anatomy, n (%)	85/95 (89.5%)
Bleeding/hemostasis, n (%)	2/95 (2.1%)
Bile duct injury, n (%)	4/95 (4.2%)
Bowel injury, n (%)	4/95 (4.2%)
Other/blank, n (%)	0/95 (0.0%)
Subtotal (bail-out) cholecystectomy, n (%)	8/726 (1.1%)
Operative time ≥120 min, n (%)	243/726 (33.5%)
Difficult cholecystectomy (composite)*, n (%)	276/726 (38.0%)
Intraoperative complication, n (%)	14/726 (1.9%)
Intraoperative complication types among complicated cases	(n=14)
Bile duct injury, n (%)	4/14 (28.6%)
Bowel injury, n (%)	9/14 (64.3%)
Bleeding/hemostasis, n (%)	1/14 (7.1%)
Other/blank, n (%)	0/14 (0.0%)
Drain use, n (%)	364/726 (50.1%)
Drain output first 24h (mL), median (IQR)	50 (50–100)
Time to oral intake (hours), median (IQR)	6 (6–6)
Length of stay (days), median (IQR)	2 (1–4)
Any postoperative complication (Clavien–Dindo ≥I), n (%)	60/726 (8.3%)
Major complication (Clavien–Dindo ≥III), n (%)	21/726 (2.9%)

*Difficult cholecystectomy was defined as conversion to open surgery and/or subtotal (bail-out) cholecystectomy and/or operative time ≥120 minutes

IQR: Interquartile range, ASA: American Society of Anesthesiologists, COPD: Chronic obstructive pulmonary disease, WBC: White blood cell, CAR: C-reactive protein-to-albumin, CRP: C-reactive protein, NLR: Neutrophil-to-lymphocyte ratio, AST: Aspartate aminotransferase, ALT: Alanine aminotransferase, GGT: Gamma-glutamyl transferase

and/or operation time ≥120 min. These three components point to the common ground [advanced inflammation, fibrosis, "frozen Calot", and failure to achieve the critical view of safety (CVS); furthermore, the current safety-focused approach explicitly emphasizes bail-out options such as conversion or subtotal when CVS cannot be achieved, changing strategies, and avoiding insistence when necessary to prevent biliary injury.[1,10] The preference for a composite definition aims to capture clinically critical decisions that might be missed by focusing only on a single endpoint such as "conversion". Current reviews highlight the heterogeneity in the definitions of DLC and the importance of reporting intraoperative endpoints (duration, conversion, complications, etc.) jointly.[6]

In our model, the choice to use logCAR rather than to categorize CAR by a cut-off value is justified by two methodological reasons: (i) CRP and ratio indices generally show a right-skewed distribution in practice; log transformation reduces the influence of outliers and provides a scale more suitable for the model assumptions; (ii) it reduces dependence on a specific threshold choice, producing a more generalizable correlation coefficient across different centers.[12]

Table 2. Comparison of non-difficult vs. difficult cholecystectomy groups (definition: conversion and/or subtotal and/or operative time ≥120 min)

Variable	Non-difficult (n=450)	Difficult (n=276)	p value
Age, years, median (IQR)	70 (67–73)	70 (67–75)	0.047
Male sex, n (%)	136/450 (30.2%)	123/276 (44.6%)	<0.001
Urgent surgery, n (%)	16/450 (3.6%)	54/276 (19.6%)	<0.001
Indication: acute cholecystitis, n (%)	12/450 (2.7%)	54/276 (19.6%)	<0.001
ASA ≥III, n (%)	63/450 (14.0%)	82/276 (29.7%)	<0.001
Anticoagulant use, n (%)	30/450 (6.7%)	43/276 (15.6%)	<0.001
WBC (×10^3/μL), median (IQR)	7.5 (6.3–9.1)	7.9 (6.5–10.2)	0.010
CRP (mg/L), median (IQR)	4.5 (1–12)	12 (4.5–45)	<0.001
Albumin (g/dL), median (IQR)	4.1 (3.8–4.4)	3.8 (3.1–4.3)	<0.001
CAR (CRP/albumin), median (IQR)	1.07 (0.26–2.86)	3.16 (1.12–12.86)	<0.001
NLR (neutrophil/lymphocyte), median (IQR)	2.17 (1.72–3.12)	2.42 (1.74–4.33)	0.003
AST (U/L), median (IQR)	22 (19–30)	23 (19–33)	0.063
ALT (U/L), median (IQR)	19 (15–26)	20 (16–28)	0.055
AST/ALT ratio, median (IQR)	1.11 (0.91–1.4)	1.11 (0.93–1.38)	0.933
GGT (U/L), median (IQR)	28 (18–44)	36 (22–78)	<0.001
Total bilirubin (mg/dL), median (IQR)	0.68 (0.5–0.9)	0.73 (0.59–1.05)	<0.001
Direct bilirubin (mg/dL), median (IQR)	0.1 (0.08–0.16)	0.13 (0.1–0.22)	<0.001
Direct/total bilirubin ratio, median (IQR)	0.17 (0.14–0.2)	0.2 (0.16–0.25)	<0.001
Operative time (min), median (IQR)	90 (74–99)	135 (120–160)	<0.001
Intraoperative complication, n (%)	0/450 (0.0%)	14/276 (5.1%)	<0.001
Length of stay (days), median (IQR)	1 (1–2)	4 (2–7)	<0.001
ICU requirement, n (%)	38/450 (8.4%)	96/276 (34.8%)	<0.001
30-day mortality, n (%)	2/450 (0.4%)	8/276 (2.9%)	0.008

IQR: Interquartile range, ASA: American Society of Anesthesiologists, WBC: White blood cell, CAR: C-reactive protein-to-albumin, CRP: C-reactive protein, NLR: Neutrophil-to-lymphocyte ratio, AST: Aspartate aminotransferase, ALT: Alanine aminotransferase, GGT: Gamma-glutamyl transferase, ICU: Intensive care unit

Table 3. Multivariable logistic regression for predictors of difficult cholecystectomy

Variable	Adjusted OR (95% CI)	p value
Age (per year)	0.98 (0.95–1.01)	0.269
Male sex	1.44 (1.02–2.03)	0.039
Urgent surgery	0.37 (0.08–1.63)	0.187
Acute cholecystitis	8.79 (2.00–38.73)	0.004
ASA ≥III	1.34 (0.85–2.12)	0.207
Anticoagulant use	1.61 (0.89–2.92)	0.117
logCAR	1.45 (1.29–1.63)	<0.001
logNLR	0.86 (0.64–1.16)	0.321
Model performance	n=726, AUC =0.727	

ASA: American Society of Anesthesiologists, OR: Odds ratio, CI: Confidence interval, AUC: Area under the curve, logCAR: log-transformed CAR, logNLR: log-transformed NLR

In our study, the difficult cholecystectomy phenotype was associated not only with prolonged operation time and "bail-out/conversion" but also with worse clinical outcomes, including increased intraoperative complications, longer length of stay, greater ICU requirement, and higher 30-day mortality. This finding suggests that, among older adults, technical difficulty may cease to be a "purely surgical issue" and instead become a risk indicator reflected in outcomes when combined with physiological fragility. Current systematic reviews and meta-analyses have shown frailty to be a significant predictor of adverse outcomes after cholecystectomy in elderly patients.[11]

The finding that acute cholecystitis, male gender, and logCAR remain independent predictors in multivariate analysis is consistent with the literature. The Tokyo approach emphasizes that inflammation in acute cholecystitis complicates anatomical dissection, and that a change in strategy should be implemented at a low threshold if safe technical steps (especially CVS) cannot be achieved.[10]

Table 4. Postoperative complications by Clavien-Dindo grade, stratified by difficulty			
Clavien–Dindo grade	**Non-difficult (n=450)**	**Difficult (n=276)**	**p value**
0	429/450 (95.3%)	221/276 (80.1%)	
I	4/450 (0.9%)	15/276 (5.4%)	
II	12/450 (2.7%)	24/276 (8.7%)	
IIIA	1/450 (0.2%)	5/276 (1.8%)	
IIIB	3/450 (0.7%)	3/276 (1.1%)	
IVA	0/450 (0.0%)	0/276 (0.0%)	
IVB	0/450 (0.0%)	0/276 (0.0%)	
V	1/450 (0.2%)	8/276 (2.9%)	
Overall p value (χ^2) for grade distribution			<0.001
Major complications (≥III) p value			<0.001

The association of male gender with surgical difficulty and conversion has also been identified as a risk factor in previous meta-analyses. Similarly, it has been reported that factors such as acute cholecystitis and age can increase the likelihood of conversion.[7,13]

The emergence of CAR as an independent predictor is biologically consistent with its composite nature, combining inflammatory burden (CRP) and physiological reserve status (albumin) under the same umbrella. Beyond acute cholecystitis, the CAR has been widely investigated as a prognostic biomarker in hospitalized older adults, in inflammatory conditions such as acute pancreatitis, and in gastrointestinal malignancies; in these contexts, elevated CAR has been associated with adverse clinical outcomes and reduced survival.[14-16] Studies have reported that CAR can predict DLC and/or conversion in series of patients with acute cholecystitis defined according to the Tokyo Guidelines 2018 criteria.[4] Although NLR can be associated with "difficulty" in some studies, its effect may not always remain independent in multivariate models because NLR may partly share the same biology as CAR on the inflammatory axis, and in the elderly population, immunosenescence, comorbidity, and drug effects may alter lymphocyte dynamics, reducing specificity. This heterogeneity is consistent with the broader evidence in the DLC literature, which shows that predictors can vary depending on center, recognition, and patient selection.[6,7,13,17]

One of our secondary findings is that laboratory parameters associated with cholestasis, especially Gamma-glutamyl transferase (GGT) and bilirubin fractions, are elevated in difficult cases. This situation can be explained by mechanisms within the clinical spectrum, such as transient obstruction accompanied by acute inflammation, microlithiasis, stone passage, or concomitant choledochal stones. However, the relationship may vary between centers, and cholestasis parameters alone may not clearly represent surgical difficulty. The association of high GGT/ALP and direct bilirubin levels with the need for conversion, reported in some studies, supports our finding.[18] On the other hand, because derived ratios such as the direct/total bilirubin ratio are not widely accepted as established predictors in the literature on DLC, we report them as exploratory findings and do not propose generalizable thresholds.

In our cohort, the aspartate aminotransferase/alanine aminotransferase (De Ritis) ratio did not differ between difficult and non-difficult cases, suggesting limited utility for predicting operative difficulty in this setting. More broadly, the evidence base for the De Ritis ratio regarding prognosis or surgical difficulty in acute cholecystitis remains limited, with heterogeneous results reported.[19] Accordingly, we present this parameter as a secondary observation rather than a primary predictor.

The fact that the vast majority of conversions occur due to "unsafe/uncertain anatomy" is consistent with current safety culture. Multi-association "safe cholecystectomy" guidelines published under the leadership of SAGES emphasize that strategies such as not insisting when CVS cannot be obtained, creating an appropriate dissection window, calling for help, and bail-out procedures (including subtotal cholecystectomy) or conversion when necessary are fundamental in preventing biliary injury.[1] Furthermore, recent studies comparing the outcomes of subtotal cholecystectomy with conversion in difficult cases show that the "bail-out" strategy should be evaluated in the context of patient selection and surgeon experience.[20] The negative impact of difficult cases on clinical outcomes in our cohort supports the value of proactive planning (senior surgeon, appropriate timing, equipment, low-threshold bail-out) in the combination of "risky patient + risky surgical site" in elderly patients.

Study Limitations

The retrospective, single-center design makes it difficult to fully control for confounders such as surgeon experience, surgical timing, details of imaging findings, and selection bias. An important characteristic of our cohort is the predominance of biliary colic cases and the relatively low proportion of acute cholecystitis. Because inflammatory burden is generally higher in acute cholecystitis, the predictive performance of CAR may vary between centers with differing proportions of acute inflammatory cases. Therefore, caution is warranted when extrapolating our findings, and external validation in cohorts enriched with acute cholecystitis is recommended. For DLC, although the composite definition is clinically significant, heterogeneity in definitions persists in the literature; this may affect external validity.[6] The lack of frailty measures (e.g., CFS) and functional outcomes in the elderly population is also a significant limitation, as frailty has been strongly shown to predict morbidity/mortality after cholecystectomy.[9] Finally, although laboratory markers reflect inflammatory biology, choledochal stones, concomitant hepatobiliary pathologies, or comorbid conditions (e.g., chronic liver disease) may affect these markers; therefore, prospective multicenter validation studies are needed.

CONCLUSION

Difficult cholecystectomy phenotype in patients ≥65 years of age is associated with clinically significant poor outcomes. The identification of CAR, using log-transformed analysis, as an independent predictor, along with acute cholecystitis and male gender, supports a practical laboratory-based approach to preoperative risk stratification. This risk stratification, when combined with safety-focused surgical strategies and low-threshold bail-out or conversion decisions, helps reduce complications in elderly patients.

Ethics

Ethics Committee Approval: Ethics approval was obtained from the University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital Ethics Committee (approval number: 2025/12-28, date: 05.01.2026). The study was conducted in accordance with the Declaration of Helsinki and its later amendments.

Informed Consent: Informed consent was waived because the study was retrospective.

Footnotes

Authorship Contributions

Surgical and Medical Practices: O.B.N., F.Di., B.G., Concept: O.B.N., F.D., Design: O.B.N., F.D., B.G., Data Collection or Processing: O.B.N., F.D., F.Di., B.G., Analysis or Interpretation: O.B.N., F.Di., B.G., Literature Search: O.B.N., F.D., F.Di., B.G., Writing: O.B.N., F.D., F.Di.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Brunt LM, Deziel DJ, Telem DA, et al; Prevention of Bile Duct Injury Consensus Work Group. Safe cholecystectomy multi-society practice guideline and state-of-the-art consensus conference on prevention of bile duct injury during cholecystectomy. Surg Endosc. 2020;34:2827-55.
2. Pisano M, Allievi N, Gurusamy K, et al. 2020 World Society of Emergency Surgery updated guidelines for the diagnosis and treatment of acute calculus cholecystitis. World J Emerg Surg. 2020;15:61.
3. Okamoto K, Suzuki K, Takada T, et al. Tokyo Guidelines 2018: flowchart for the management of acute cholecystitis. J Hepatobiliary Pancreat Sci. 2018;25:55-72. Erratum in: J Hepatobiliary Pancreat Sci. 2019;26:534.
4. Utsumi M, Sakurai Y, Narusaka T, et al. C-reactive protein to albumin ratio predicts difficult laparoscopic cholecystectomy in patients with acute cholecystitis diagnosed according to the Tokyo Guidelines 2018. Asian J Endosc Surg. 2022;15:487-94.
5. Lee SK, Lee SC, Park JW, Kim SJ. The utility of the preoperative neutrophil-to-lymphocyte ratio in predicting severe cholecystitis: a retrospective cohort study. BMC Surg. 2014;14:100.
6. Abdallah HS, Sedky MH, Sedky ZH. The difficult laparoscopic cholecystectomy: a narrative review. BMC Surg. 20250;25:156.
7. Magnano San Lio R, Barchitta M, Maugeri A, Quartarone S, Basile G, Agodi A. Preoperative risk factors for conversion from laparoscopic to open cholecystectomy: a systematic review and meta-analysis. Int J Environ Res Public Health. 2022;20:408.
8. Dindo D, Demartines N, Clavien PA. Classification of surgical complications: a new proposal with evaluation in a cohort of 6336 patients and results of a survey. Ann Surg. 2004;240:205-13.
9. Clavien PA, Barkun J, de Oliveira ML, et al. The Clavien-Dindo classification of surgical complications: five-year experience. Ann Surg. 2009;250:187-96.
10. Wakabayashi G, Iwashita Y, Hibi T, et al. Tokyo Guidelines 2018: surgical management of acute cholecystitis: safe steps in laparoscopic cholecystectomy for acute cholecystitis (with videos). J Hepatobiliary Pancreat Sci. 2018;25:73-86.
11. Niknami M, Tahmasbi H, Firouzabadi SR, et al. Frailty as a predictor of mortality and morbidity after cholecystectomy: a systematic review and meta-analysis of cohort studies. Langenbecks Arch Surg. 2024;409:352.
12. Grund B, Sabin C. Analysis of biomarker data: logs, odds ratios, and receiver operating characteristic curves. Curr Opin HIV AIDS. 2010;5:473-9.
13. Wevers KP, van Westreenen HL, Patijn GA. Laparoscopic cholecystectomy in acute cholecystitis: C-reactive protein level combined with age predicts conversion. Surg Laparosc Endosc Percutan Tech. 2013;23:163-6.
14. Ranzani OT, Zampieri FG, Forte DN, Azevedo LC, Park M. C-reactive protein/albumin ratio predicts 90-day mortality of septic patients. PLoS One. 2013;8:e59321.

15. Kaplan M, Ates I, Akpinar MY, et al. Predictive value of C-reactive protein/albumin ratio in acute pancreatitis. Hepatobiliary Pancreat Dis Int. 2017;16:424-30.
16. Kinoshita A, Onoda H, Imai N, et al. The C-reactive protein/albumin ratio, a novel inflammation-based prognostic score, predicts outcomes in patients with hepatocellular carcinoma. Ann Surg Oncol. 2015;22:803-10.
17. Wei HH, Wang YX, Xu B, Zhang YG. Preoperative systemic and local inflammation are independent risk factors for difficult laparoscopic cholecystectomy after percutaneous transhepatic gallbladder drainage. Heliyon. 2024;10:e36081.
18. Özdemir A, Karakaya A, Pergel A. Conversion from laparoscopic cholecystectomy to open surgery: reasons and possible risks: a single center experience. Deneysel ve Klinik Tip Dergisi. 2022;39:781-5.
19. Gürlevik E, Akça HŞ, Akça MT, Akça H. The role of the De Ritis ratio in acute cholecystitis: a retrospective observational study. J Contemp Med. 2023;13:388-95.
20. Ramírez-Giraldo C, Monroy DC, Isaza-Restrepo A, et al. Subtotal laparoscopic cholecystectomy versus conversion to open as a bailout procedure: a cohort study. Surg Endosc. 2024;38:4965-75.

DOI: 10.4274/forbes.galenos.2026.37167
Forbes J Med 2026;7:23-32

Original Article

Contemporary Outcomes of Chorionic Villus Sampling from the First Trimester to Neonatal Follow-Up

Zübeyde EMİRALİOĞLU ÇAKIR[1], Hakan GÖLBAŞI[2], Burak BAYRAKTAR[3], Selviye HALKSEVER ÖZVATAN[4], Raziye TORUN[1], İlayda GERCİK ARZIK[2], Hale ANKARA AKTAŞ[2], Emre BAYRAM[2], Selcan KESAN[5], Altuğ KOÇ[5], Atalay EKİN[2]

[1]University of Health Sciences Türkiye, İzmir City Hospital, Division of Perinatology, Clinic of Obstetrics and Gynecology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Perinatology, İzmir, Türkiye
[3]University of Health Sciences Türkiye, Ankara Etlik City Hospital, Clinic of Perinatology, Ankara, Türkiye
[4]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Obstetrics and Gynecology, İzmir, Türkiye
[5]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Medical Genetics, İzmir, Türkiye

Cite as: Emiralioğlu Çakır Z, Gölbaşı H, Bayraktar H, et al. Contemporary outcomes of chorionic villus sampling from the first trimester to neonatal follow-up. Forbes J Med. 2026;7:23-32

ABSTRACT

Objective: To evaluate contemporary indications, diagnostic yield, diagnostic limitations, and pregnancy–neonatal outcomes of chorionic villus sampling (CVS) in a tertiary referral population in the era of widespread non-invasive prenatal testing.

Methods: This retrospective cohort study included 70 singleton pregnancies that underwent CVS at a tertiary referral center between October 2023 and June 2025. Maternal characteristics, indications for CVS, genetic testing results, and procedure-related outcomes were recorded. Pregnancy loss, the need for repeat invasive testing, termination decisions, and delivery and neonatal outcomes were assessed.

Results: Major fetal structural anomalies were the leading indication (40%). Pathological genetic findings were identified in 21.4% of cases. Despite multimodal testing, 5.7% remained without a result, and 17.1% required repeat invasive sampling. No immediate complications occurred. Four pregnancy losses before 24 weeks' gestation (5.7%) were observed. Three occurred in pregnancies with major structural or chromosomal abnormalities, while one occurred in a structurally normal fetus with inconclusive cytogenetic results. When anomaly-associated cases were excluded, the observed loss rate among structurally normal pregnancies was 1.4%. No membrane rupture or chorioamnionitis occurred. Among live births, 97.2% were delivered at term. Neonatal outcomes were reassuring, and stillbirths (3.1%) were attributable to severe fetal or maternal pathology rather than the CVS procedure.

Conclusion: Structural fetal anomalies are now the leading indication for CVS. Post-procedure complications were rare, and most adverse outcomes were observed in pregnancies with underlying fetal or maternal pathology. These findings support the continued role of CVS as a diagnostic option in selected high-risk pregnancies.

Keywords: Chorionic villus sampling, prenatal diagnosis, genetic testing, chromosomal abnormalities, pregnancy outcomes, diagnostic yield, non-invasive prenatal testing, structural anomalies

Received: 31.12.2025
Accepted: 19.03.2026
Epub: 23.03.2026
Publication Date: 02.04.2026

Corresponding Author:
Zübeyde EMİRALİOĞLU ÇAKIR
University of Health Sciences Türkiye, İzmir City Hospital, Division of Perinatology, Clinic of Obstetrics and Gynecology, İzmir, Türkiye
✉ zubeydeemiralioglu@hotmail.com
ORCID ID: 0000-0001-8298-8363

INTRODUCTION

Prenatal identification of genetic and chromosomal abnormalities has become a central component of contemporary obstetric and perinatal care. The detection of numerical and structural chromosomal anomalies, as well as of single-gene disorders, plays a critical role in risk assessment, counseling, and early pregnancy management.[1] Invasive diagnostic procedures such as chorionic villus sampling (CVS) and amniocentesis (AC) remain essential tools in this process and are supported by current professional guidelines, including those of the American College of Obstetricians and Gynecologists (ACOG), the Royal College of Obstetricians and Gynaecologists, and the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG).[2-4]

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

CVS is typically performed between 10 and 14 weeks of gestation and enables retrieval of fetal genetic material several weeks earlier than AC, which is generally not recommended before 15 weeks because of an increased risk of complications.[1,5] Achieving a diagnosis in the first trimester provides several important advantages, including reduced parental anxiety, earlier clinical decision-making, and the opportunity to perform pregnancy termination at a safer gestational age (GA) when indicated.[1,6,7] Additional procedural benefits include the absence of direct fetal manipulation and of disruption to the fetal membranes. For these reasons, CVS has long been regarded as a reliable and effective first-trimester diagnostic option in experienced centers.[8]

Despite these strengths, several technical limitations may influence diagnostic accuracy. Maternal cell contamination (MCC), confined placental mosaicism, and culture failure can complicate the interpretation of CVS results and may necessitate further invasive testing in a subset of patients.[9,10] Reported culture failure rates following CVS range from approximately 2.2% to 34%, while MCC rates range from 1.06% and 24.6%, reflecting substantial variability across centers and laboratory methodologies.

Although recent advances in prenatal screening, particularly non-invasive prenatal testing (NIPT), have reduced the number of invasive procedures, accurate counseling remains essential when offering CVS.[11,12] Reported procedure-related pregnancy loss rates range from 0.7% to 3.2%.[1,9,12,13] This variability may reflect differences in loss definitions, GA thresholds, and the inclusion of structurally or chromosomally abnormal pregnancies. Accordingly, counseling should incorporate accurate center-specific outcome data.[12]

The aim of this study was to evaluate the indications, diagnostic efficiency, procedural outcomes, and short-term pregnancy and neonatal outcomes of CVS performed in a tertiary referral center. By examining genetic testing methods, post-procedural complications, the need for repeat invasive testing, and subsequent pregnancy and neonatal outcomes, this study aims to provide contemporary outcome data from a high-risk referral population and to help inform clinical counseling and decision-making.

METHODS

This retrospective cohort included singleton pregnancies that underwent CVS between October 2023 and June 2025 at the Maternal–Fetal Medicine Unit of the University of Health Sciences Türkiye, İzmir City Hospital. The study protocol was approved by the University of Health Sciences Türkiye, İzmir City Hospital Ethics Committee (approval number: 2025/325, date: 09.07.2025).

Study Population and Indications

Pregnancies in which CVS was performed for genetic diagnosis were included in the study. Maternal age, gravidity, parity, number of living children, and GA at the time of the procedure were recorded. GA was determined by the last menstrual period or by first-trimester crown–rump length. Indications for CVS were categorized as follows: maternal anxiety, high-risk combined first-trimester screening test (cut-off >1/250), positive NIPT result, sonographic soft markers, major fetal structural anomalies, history of a previous child with aneuploidy or a genetic disease, parental carrier status for a chromosomal abnormality, and multiple indications (defined as the presence of more than one indication in the same patient).

Nuchal translucency (NT), included in the sonographic marker group, was measured by a maternal–fetal medicine specialist in accordance with the guidelines of the American Institute of Ultrasound in Medicine. Increased NT was defined as a measurement ≥3.0 mm during first-trimester sonography.[14]

All patients underwent a standardized first-trimester ultrasound examination in accordance with the ISUOG guidelines.[15] All ultrasound examinations were performed using a Voluson E8 system (GE Healthcare, Wauwatosa, WI, USA) equipped with a 2–9 MHz convex transducer.

CVS Technique

All procedures were performed between 11 and 14 weeks of gestation. Prior to CVS, detailed counseling was provided about the results of screening tests, available invasive diagnostic options, and the benefits and limitations of each option. This was a retrospective study, and written informed consent was obtained from all patients.

Fetal viability, presentation, and placental location were evaluated before the procedure. CVS was performed under continuous ultrasound guidance using a double-needle technique, with the same ultrasound system and standardized protocol used for the first-trimester screening examinations. An 18-gauge needle was introduced transabdominally, without local anesthesia, to reach the placenta. After removal of the stylet, an assistant inserted a 20-gauge aspiration needle and manually aspirated placental villi. All procedures were carried out under aseptic conditions by a fellow physician under the direct supervision of a maternal–fetal medicine specialist. Immediately after the procedure, fetal heart activity, amniotic fluid volume, and the presence of any hemorrhage were assessed. Rh-negative and indirect Coombs-negative patients received 300 µg of anti-D immunoglobulin for prophylaxis against alloimmunization.

Genetic Testing Protocol

All specimens were processed in the institutional genetics laboratory. Quantitative fluorescence polymerase chain reaction (QF-PCR) and conventional karyotyping were routinely performed.

Chromosomal microarray analysis (CMA) is not routinely performed in our clinic; instead, it is reserved for pregnancies with major fetal structural anomalies or increased NT, as well as for cases in which conventional karyotyping yields no result or yields a normal result despite high clinical suspicion. Next-generation sequencing (NGS) or Multiplex Ligation-dependent Probe Amplification (MLPA) was performed in a limited number of cases, including pregnancies with a relevant family history or suspicion of a single-gene disorder. The total number of patients evaluated with each genetic method was recorded.

Patients with abnormal results on any genetic test were assigned to the pathological group. Those in whom no results could be obtained from any genetic test were categorized as the no-result group. The number of pregnancy terminations within each indication group was also recorded.

Maternal, Pregnancy, and Neonatal Outcomes

Maternal, pregnancy, and neonatal outcomes included complications and pregnancy loss, such as vaginal bleeding, preterm premature rupture of membranes (PPROM), chorioamnionitis, and miscarriage. To provide more detailed information and to address potential discrepancies in definitions, we categorized all spontaneous losses into the following time intervals: 0–48 hours, 2–7 days, 7–14 days, and 2–24 weeks post-procedure. Patients who required repeat invasive testing, as well as those who accepted or declined repeat sampling, were recorded. Additional follow-up data included counseling on pregnancy termination; the number of terminations performed; maternal or fetal complications; hospital admissions and their indications; and stillbirths and live-birth outcomes. For subsequent pregnancies following CVS, follow-up data were collected from both hospital records and structured telephone interviews to investigate additional complications not documented in medical records. Birth characteristics, including preterm or term delivery, mode of delivery, and neonatal outcomes, were recorded for all live births. However, third-trimester follow-up data were unavailable for six patients. Therefore, variables related to late-pregnancy follow-up, hospital admissions, associated maternal–fetal conditions, and delivery and neonatal outcomes were evaluated in the remaining 64 patients.

Outcome Definitions

Primary outcomes were the diagnostic yield of CVS (proportion of pathological genetic findings across QF-PCR, karyotyping, CMA, NGS, and MLPA) and pregnancy loss rates (miscarriage, intrauterine death, stillbirth) from the procedure to neonatal follow-up.

Secondary outcomes included post-procedure complications (vaginal bleeding, PPROM, chorioamnionitis); need for repeat invasive testing; pregnancy termination rates; maternal/fetal complications [e.g., preeclampsia, fetal growth restriction (FGR)]; delivery characteristics (preterm birth, mode of delivery); and neonatal outcomes [birth weight, Apgar scores, neonatal intensive care unit (NICU) admission, neonatal loss].

In the present study, procedure-related pregnancy loss was not defined using a strict causal attribution model. Instead, pregnancy losses were categorized descriptively by the time interval following CVS and by the presence or absence of major fetal structural or chromosomal abnormalities.

Statistical Analysis

Descriptive statistical analyses were performed using the Statistical Package for the Social Sciences, version 26.0 (IBM Corp., Armonk, NY). Continuous variables were summarized as mean ± standard deviation, and as ranges when appropriate, while categorical variables were reported as frequencies and percentages. Exact binomial 95% confidence intervals (CIs) were calculated for key proportions. Exploratory subgroup comparisons were performed using Fisher's exact test. Univariable logistic regression analysis was conducted to evaluate predictors of pathological genetic findings.

RESULTS

A total of 80 patients were initially scheduled for CVS between 2023 and 2025. The procedure could not be performed in 10 patients (12.5%): nine due to an unfavorable placental location (posterior; 11.3%) and one due to a pre-procedural missed miscarriage (1.3%), resulting in a final study cohort of 70 patients. A total of seven patients who were Rh-negative and indirect Coombs-negative received 300 µg of anti-D immunoglobulin. The mean maternal age was 32.2±5.8 years, and CVS was performed at a mean GA of 12.9±0.8 weeks. Demographic characteristics of the study population are summarized in Table 1.

Table 2 presents the indications for CVS and associated diagnostic outcomes. Major fetal structural anomalies were the most frequent indication (40%, 28/70), among which there were 11 pathological results (39.3%) and 16 terminations (57.1%). An increased risk on the first-

trimester combined screening test accounted for 22.9% of cases; none of these cases yielded pathological genetic results. Maternal anxiety was the indication in 3 cases (4.3%), whereas an increased risk on NIPT (2.9%) resulted in abnormal findings in both patients. Sonographic soft markers represented 10% of indications, most commonly increased NT. One pathological result (14.3%) was identified in this group.

A prior pregnancy affected by a chromosomal abnormality was the indication for 6 patients (8.6%); none of these pregnancies showed a pathological finding, although one pregnancy was later terminated after micromelia suggestive of a skeletal dysplasia. Parental carrier status accounted for 2 cases (2.9%); one pathological result and one inconclusive result. Multiple indications were present in 6 patients (8.6%), including one patient with a pathogenic neurofibromatosis type 1 (NF1) variant. In another case, MCC was inconclusive, and the patient subsequently underwent AC, which identified a homozygous osteogenesis imperfecta-related variant, leading to pregnancy termination. Exploratory subgroup comparisons of diagnostic yield across indication groups were performed using Fisher's exact test. The rate of pathological findings was significantly higher in pregnancies with major fetal structural anomalies than in those with other indications ($p<0.001$). Univariable logistic regression analysis demonstrated that the presence of major fetal structural anomalies was significantly associated with pathological genetic findings (odds ratio 10.7%; 95% CI: 3.1–36.2; $p<0.001$). Overall, pathological genetic findings were identified in 15 of 70 cases (21.4%; 95% CI: 12.8–32.3). No diagnostic result was obtained in 4 of 70 cases (5.7%; 95% CI: 1.6–13.9), and pregnancy termination was performed in 20 of 70 cases (28.6%; 95% CI: 18.4–40.6).

Table 1. Demographic characteristics of the study population

n=70	Mean ± SD (min–max)
Age	32.2±5.8 (20–43)
Gravidity	2.3±1.4 (1–7)
Parity	0.94±0.98 (0–4)
Living children	0.90±0.97 (0–4)
Abortions	0.40±0.97 (0–5)
Gestational age (weeks)	12.9±0.8 (11–14)

SD: Standard deviation, min: Minimum, max: Maximum

Table 3 summarizes the genetic test results. QF-PCR identified 9 pathological findings (12.9%), including 45, X (monosomy X; Turner syndrome), trisomy 21 (n=3), trisomy 18 (n=3), and trisomy 13 (n=2), while 14 samples were non-informative. Conventional karyotyping detected 12 abnormalities (17.1%), including del(6)(p22), monosomy X, mosaic Turner syndrome, and multiple autosomal trisomies. Nineteen samples were inconclusive due to low band resolution, culture failure, MCC, or technical issues.

CMA was performed in 13 cases and identified three abnormalities (23.1%): a pathogenic 6p25.3–p24.1 microdeletion, trisomy 21, and monosomy X. The latter two had already been detected by karyotyping; thus, the 6p25.3–p24.1 microdeletion represented the additional diagnostic yield of CMA beyond conventional karyotyping (1/13 cases; 7.7%). NGS identified two likely pathogenic variants (NF1 and COL1A1), while one sample could not be analyzed due to MCC. MLPA confirmed one homozygous

Table 2. Clinical indications for CVS and associated diagnostic outcomes

n=70	n (% of total)	Pathological result n (% within group)	No result n (% within group)	TOP n (% within group)
Maternal anxiety	3 (4.3)	-	1 (33.3)	-
Screen-positive test (1st trimester)	16 (22.9)	-	-	-
Increased risk in NIPT	2 (2.9)	2 (100)	-	2 (100)
Sonographic markers	7 (10)	-	1 (14.3)	-
Increased NT	5 (7.1)	-	-	-
Others	2 (2.9)	-	1 (50)	-
Major fetal structural anomaly	28 (40)	11 (39.3)	1 (3.6)	16 (57.1)
Previous pregnancy with chromosomal abnormality	6 (8.6)	-	-	1 (16.7)
Parental chromosomal rearrangement carrier	2 (2.9)	1 (50)	1 (50)	1 (50)
Multiple indications	6 (8.6)	1 (16.7)	-	1 (16.7)
Total	70 (100)	15 (21.4)	4 (5.7)	20 (28.6)

TOP: Termination of pregnancy, NIPT: Non-invasive prenatal testing, NT: Nuchal translucency, CVS: Chorionic villus sampling

Table 3. Genetic test results

Type of test n=70	n (% of total)	Pathological result n (%)	Normal result n (%)	No result n (%)
QF-PCR	70 (100)	9 (12.9) 1 45,X 3 T21 3 T18 2 T13	47 (67.1)	14 (20)[a]
KT	70 (100)	12 (17.1) 1 del(6)(p22) 1 45,X 1 47,XXX/45,X/46,XX (mosaic Turner syndrome) 3 T21 3 T18 3 T13	39 (55.7)	19 (27.1)[b]
CMA	13 (18.6)	3 (23.1) 1 6p25.3–p24.1 del (pathogenic) 1 T21 1 45,X	8 (61.5)	2 (15.4)[c]
NGS	4 (5.7)	2 (50) 1 NF1 c.1466A>G (p.Tyr489Cys) 1 COL1A1 p.Gly626Asp (heterozygous, likely pathogenic)	1 (25)	1 (25)
MLPA	1 (1.4)	1 (100) SMN1 exons 7–8 (homozygous deletion)		

[a]In 8 cases, no result was obtained due to maternal cell contamination. It was recommended to wait for the cytogenetic results for 2 patients with chromosome 13 anomaly, 1 patient with chromosome 18 anomaly, and 2 patients with sex chromosomes

[b]No result due to low band resolution (n=3), maternal cell contamination (n=3), culture failure (n=7), or technical issues (n=6)

[c]No result due to maternal cell contamination (n=1) or technical issues (n=2)

Maternal cell contamination was observed in 20 cases overall and contributed to inconclusive results in QF-PCR (n=4), karyotyping (n=2), QF-PCR + karyotyping (n=2), and QF-PCR + CMA (n=2). For each test modality, percentages in the "Pathological," "Normal," and "No result" columns reflect within-group distributions, calculated relative to the number of patients who underwent that specific test ("n tested").

QF-PCR: Quantitative fluorescence polymerase chain reaction, CMA: Chromosomal microarray analysis, KT: Conventional karyotype analysis, NGS: Next-generation sequencing, MLPA: Multiplex ligation-dependent probe amplification, T: Trisomy

Note: For each test modality, percentages in the "pathological," "normal," and "no result" columns reflect within-group distributions, calculated relative to the number of patients who underwent that specific test ("n tested")

SMN1 deletion. Overall, pathological findings in the cohort included trisomy 21 (4.3%), trisomy 18 (4.3%), trisomy 13 (4.3%), monosomy X (1.4%), mosaic Turner syndrome (1.4%), a pathogenic 6p deletion (1.4%), an NF1 variant (1.4%), a homozygous SMN1 deletion (1.4%), and a likely pathogenic COL1A1 variant (1.4%).

A total of 70 CVS procedures were evaluated for short- and intermediate-term outcomes (Table 4). No immediate complications occurred within the first 48 hours. Two patients experienced early pregnancy loss within the first two weeks (2.8%; 95% CI: 0.3–9.7). Repeat invasive testing was required in 12 patients (17.1%) because initial results were inconclusive or incomplete, most commonly secondary to MCC, culture failure, or other technical limitations.

Termination of pregnancy was recommended in 23 cases (32.9%) and performed in 20 cases (28.6%). The mean GA at termination was 15.4±3.3 weeks (range: 11–22). Additional maternal–fetal complications occurred in two pregnancies (3.1%): one pregnancy developed preeclampsia with placental abruption, and the other involved a fetus with trisomy 21 that demonstrated early-onset FGR. Hospitalization was required in three cases (4.7%): one case due to FGR related to trisomy 21 and two cases due to complications associated with diabetes.

Stillbirth occurred in two pregnancies (3.1%): one associated with multiple fetal anomalies and the other in the setting of severe preeclampsia with placental abruption. Among the 64 pregnancies with available delivery outcomes, 36 resulted in live births, in line with expectations for a clinically high-risk CVS population.

Pregnancy loss between 2–24 weeks occurred in two additional cases (2.9%), both were associated with major fetal anomalies: megacystis with trisomy 13 and cystic hygroma with trisomy 18. The detailed clinical characteristics of all pregnancy loss cases are summarized in Table 5.

Table 4. Pregnancy outcomes after CVS		
	Mean ± SD or n (%)	**Description/notes**
Early and ıntermediate outcomes (n=70)		
Complication/loss (0–2 days)	0 (0)	
Complication/loss (2–7 days)	1 (1.4)	
Complication/loss (7–14 days)	1 (1.4)	
Complication/loss (2–24 weeks)	2 (2.8)	
Repeat invasive procedure required	12 (17.1)	5/12 (41.7%) accepted repeat testing*
Termination recommended	23 (32.9)	
Termination performed	20 (28.6)	
Gestational age at termination (weeks)	15.4±3.3 (range 11–22)	
Late pregnancy and delivery outcomes (n=64)		
Maternal/fetal complications	2 (3.1)	1 severe preeclampsia with placental abruption 1 early-onset FGR associated with T21
Hospitalization required	3 (4.7)	1 early-onset FGR associated with T21 2 preterm labor
Stillbirth	2 (3.1)	1 multiple anomalies 1 preeclampsia with abruption
Live birth	36 (56.3)	

*Structural evaluation was inconclusive due to maternal cell contamination (n=2), low band resolution (n=3), culture failure (n=6), or need for additional genetic investigations (n=1)

FGR: Fetal growth restriction, SD: Standard deviation, CVS: Chorionic villus sampling, T: Trisomy

Table 5. Detailed characteristics of pregnancy loss cases following CVS

Case	Maternal age	GA at CVS (weeks)	GA at pregnancy loss (weeks)	Structural anomaly	Genetic result	MCC	Culture failure	Clinical context
1	42	13	14	-	QF-PCR normal; culture failure	-	1	Screen positive test results
2	35	12	14	Megacystis, increased NT	T13	-	-	Megacystis, increased NT
3	31	13	13	Megacystis	Normal	1	-	Megacystis
4	31	11	14	Cystic hygroma	Normal	-	-	Cystic hygroma

GA: Gestational age, CVS: Chorionic villus sampling, MCC: Maternal cell contamination, QF-PCR: Quantitative fluorescence polymerase chain reaction, NT: Nuchal translucency

Overall, pregnancy loss occurred in 4 of 70 pregnancies (5.7%). Three losses were associated with major fetal structural or chromosomal abnormalities, whereas one occurred in a structurally normal fetus with inconclusive cytogenetic results. When pregnancies with major anomalies were excluded, the observed loss rate among structurally normal pregnancies was 1.4% (1/70).

Delivery outcomes among the 36 live births are summarized in Table 6. Of these, 35 (97.2%) delivered at term, while one patient (2.8%) delivered preterm at 35 weeks due to trisomy 21. Spontaneous vaginal delivery occurred in 25% of cases, and primary cesarean section occurred in 30.6% of cases. One neonatal death (2.8%) occurred due to cardiac failure related to trisomy 21. The mean GA at delivery was 38.2±1.1 weeks, and the mean birth weight was 3131±570 g. NICU admission was required for 13.9% of neonates; sex distribution was balanced (52.8% female and 47.2% male).

DISCUSSION

This study provides a detailed evaluation of CVS indications, diagnostic yield, and associated pregnancy and neonatal outcomes in this cohort. The results

Table 6. Delivery and neonatal outcomes	
n=36	**Mean ± SD or n (%)**
Term delivery	35 (97.2)
Preterm delivery (<37 weeks)	1 (2.8)*
Spontaneous vaginal delivery	9 (25)
Primary cesarean section	11 (30.6)
Repeat cesarean section	16 (44.4)
Neonatal loss	1 (2.8)**
Gestational age at birth (weeks)	38.2±1.1 (range 35–41)
Birth weight (g)	3131.3±570.1
Female newborns	19 (52.8)
Male newborns	17 (47.2)
Apgar score (1 min)	7.89±1.01
Apgar score (5 min)	9.03±0.97
NICU admission	5 (13.9)

*35 weeks, T21

**Neonatal death secondary to cardiac failure, T21

g: Grams, min: Minute, NICU: Neonatal intensive care unit, SD: Standard deviation, T: Trisomy

demonstrate that major fetal structural anomalies represent the leading indication for CVS and are associated with the highest rate of pathological genetic findings. This association was further supported by exploratory statistical analyses, including Fisher's exact test and univariable logistic regression, which showed a significantly higher likelihood of pathological findings in pregnancies with structural anomalies. In contrast, post-procedure pregnancy loss rates were low in this cohort and occurred predominantly in pregnancies with significant underlying fetal or maternal risk factors.

Pathological genetic results were identified in 21.4% of cases. Although trisomy 21 is typically the most commonly reported aneuploidy, trisomy 18 occurred at a similar frequency in our cohort.[7,9,16] This pattern likely reflects the high prevalence of structural anomalies in our population (40%), which exceeds rates reported in most studies, in which an increased risk on first-trimester screening tests is usually the most common indication.[1,9,12] This may be related to the profile of a tertiary referral center with advanced ultrasound technology and experienced fetal imaging specialists. Moreover, the widespread adoption of NIPT in recent years has not only begun to replace traditional first-trimester screening but has also led to a reduction in invasive diagnostic procedures among women identified as high-risk by first-trimester screening.[17,18] This shift may also explain why elevated first-trimester screening risk is no longer the most common indication for CVS in our cohort. Current guidelines indicate that NIPT is a screening test requiring confirmation by invasive procedures such as CVS or AC.[19] Nevertheless, rare autosomal trisomies are frequently associated with fetal or confined placental mosaicism, which may result in discordance between placental and fetal genetic findings; therefore, AC may be preferred to confirm true fetal involvement when mosaicism is suspected.[20,21]

Changes in clinical practice have also affected the role of advanced maternal age (AMA). Current guidelines from ACOG no longer recommend AMA as a stand-alone indication for invasive testing.[22] Nevertheless, many published series still report that CVS procedures were performed solely for AMA.[1,9,23] In our cohort, no patients underwent CVS solely for AMA. Relatively few studies appear to reflect this recent shift in clinical practice.[12] Spinal muscular atrophy (SMA) carrier screening is now widely recommended.[24,25] Several studies have supported the use of CVS for prenatal diagnosis in pregnancies at risk for SMA.[26-29]

Although our sample size was limited, our findings highlight the relevance of CVS for SMA carrier couples. As carrier screening programs expand, prenatally diagnosed SMA cases are expected to increase, further strengthening the evidence base for this indication.

Reported pregnancy loss rates after CVS vary across studies, largely due to differences in study design and definitions of procedure-related pregnancy loss.[1,9,11-13] Previous studies have reported higher pregnancy loss rates in women aged ≥40 years, both following CVS and in the general obstetric population.[30] In addition, several studies have evaluated both the expected miscarriage risk based on maternal and fetal characteristics and the observed miscarriage rates following CVS.[12,31] In these analyses, no significant increase in miscarriage risk attributable to the procedure was observed. In our study, pregnancy losses were reported according to commonly used time-based categories. No immediate complications, such as vaginal bleeding or PPROM, were observed; however, four pregnancy losses occurred before 20 weeks. Three were associated with major structural or chromosomal abnormalities, while one occurred in a pregnancy without identifiable structural or chromosomal abnormalities. When anomaly-associated cases were excluded, the observed loss rate among structurally normal pregnancies was 1.4%. Pregnancy losses observed following CVS appeared to be predominantly attributable to underlying fetal pathology or AMA. Post-procedure loss rates remained low in this cohort. However, establishing any causal link to the procedure is difficult because of the retrospective design, small sample size, low number of events, and confounding by indication in this high-risk population.

CVS remains an important option when confirmatory testing is required for pregnancy termination decisions, despite advances in NIPT.[32] Published termination rates following CVS range from 16% to 29%, and our findings fall within the upper range of previously reported rates.[1,9,12] This finding likely reflects early and accurate prognostic assessment enabled by detailed ultrasound evaluation and contemporary genetic testing strategies.

Non-informative results can occur due to confined placental mosaicism, insufficient villus sampling, or limited cytogenetic resolution.[1,10] When considered individually, culture failures were observed in 27.1% of cases; however, only 5.7% of patients remained without a definitive diagnosis after applying additional genetic testing methods. These findings are consistent with the potential benefit of multimodal testing strategies. Nonetheless, 17.1% of patients required repeat invasive testing, and more than half declined a second procedure, highlighting the need to consider AC early, particularly when mosaicism is suspected.[9] MCC was an important cause of inconclusive genetic results in our cohort. Although MCC is a recognized limitation of CVS, the occurrence of MCC may also reflect sampling technique, placental characteristics, or laboratory processing factors. Therefore, reducing MCC rates through optimized sampling strategies and laboratory protocols may improve diagnostic efficiency and decrease the need for repeat invasive procedures. However, the determinants of CVS sampling failure remain incompletely understood.[10] Chevreau et al.[10] reported a 34% failure rate, independent of operator experience, and emphasized the need for careful pre-procedural assessment to reduce unsuccessful sampling and repeat procedures.

Regarding obstetric complications, Movahedi et al.[33] reported that PPROM and FGR were the most frequent observed complications following CVS, while no cases of chorioamnionitis were identified. In our cohort, no cases of PPROM were observed, and FGR occurred in only one pregnancy affected by trisomy 21. Similarly, other studies have reported chorioamnionitis as a rare complication following CVS.[33,34]

In our cohort, preeclampsia occurred in only one of 44 ongoing pregnancies that had third-trimester follow-up. This case was complicated by placental abruption and stillbirth. Although the overall prevalence was comparable to that reported in the general population (4.6%),[35] the severity of this case highlights the importance of close monitoring.

Neonatal outcomes following CVS have been evaluated in a limited number of studies.[36] In that large retrospective cohort, the mean GA at delivery was 38.4 weeks, with a preterm birth rate of 10.5%. Cesarean delivery occurred in 53.0% of cases; the mean birth weight was approximately 3158 g, and NICU admission was required in 13.0% of neonates; there was no significant increase in composite neonatal morbidity compared with AC. Our study provides a comprehensive evaluation of indications, genetic findings, and perinatal, obstetric, and neonatal outcomes within a single cohort. Despite the high-risk characteristics of the population, delivery outcomes were generally favorable. Most pregnancies reached term (mean GA of 38.2 weeks); preterm delivery occurred in only one case (2.8%; trisomy 21), and neonatal loss was limited to a single case of trisomy 21. Stillbirth (3.1%) was attributable to severe fetal or maternal pathology rather than to the CVS procedure. The high prevalence of previous cesarean deliveries reflected the underlying obstetric risk profile. Overall, obstetric and neonatal outcomes in this tertiary referral cohort were generally favorable, although these observations are tempered by the study's limitations.

The present study has several limitations. First, the relatively small sample size and the limited number of events reduce the statistical power of the analyses and preclude more extensive multivariable modeling. Second, the retrospective single-center design introduces the potential for referral bias, because our institution is a tertiary referral center that frequently receives high-risk pregnancies with suspected fetal anomalies. Third, confounding by indication cannot be excluded, since pregnancies undergoing CVS often have underlying clinical risk factors that may independently influence pregnancy outcomes. Fourth, no predefined adjudication criteria were applied to attribute pregnancy loss specifically to the CVS procedure rather than to underlying fetal or maternal pathology; therefore, causal attribution should be interpreted with caution. Although all procedures were performed by fellows under the supervision of maternal–fetal medicine specialists, operator experience may influence procedural success, MCC, and complication rates, and should therefore be considered when interpreting the results. Finally, long-term neonatal or neurodevelopmental follow-up data were not available, limiting the ability to assess potential late outcomes beyond the perinatal period.

CONCLUSION

These findings suggest that the indications for CVS are increasingly shifting toward pregnancies with structural fetal anomalies. Our findings indicate that CVS continues to play an important role in definitive prenatal diagnosis, accurate counseling, and timely pregnancy management and that observed post-procedure complication rates were low in both the prenatal and neonatal periods. Post-procedure pregnancy losses in this cohort predominantly occurred in pregnancies with significant fetal or maternal

risk factors and were less frequently observed in structurally normal cases. The results also emphasize the importance of informing families that, in selected cases, additional invasive testing may be required. However, given the retrospective design and single-center nature of the study, these observations should be interpreted cautiously and confirmed in larger, multicenter prospective studies to enhance generalizability and strengthen the evidence base for clinical practice.

Ethics

Ethics Committee Approval: The study protocol was approved by the University of Health Sciences Türkiye, İzmir City Hospital Ethics Committee (approval number: 2025/325, date: 09.07.2025).

Informed Consent: This was a retrospective study, and written informed consent was obtained from all patients.

Acknowledgments

The authors would like to thank the Maternal–Fetal Medicine Unit and the Medical Genetics Department of University of Health Sciences Türkiye, İzmir City Hospital for their collaboration and support throughout the study. We also extend our appreciation to the nursing staff for their contributions to patient care and follow-up.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Z.E.Ç., H.G., B.B., İ.G.A., H.A.A., E.B., S.K., A.K., A.E., Concept: Z.E.Ç., H.G., A.E., Design: Z.E.Ç., H.G., Data Collection or Processing: S.H.Ö., R.T., İ.G.A., H.A.A., E.B., Analysis or Interpretation: Z.E.Ç., S.H.Ö., Literature Search: Z.E.Ç., H.G., Writing: Z.E.Ç., H.G., A.E.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Öztürk FH, Öcal FD, Erol SA, et al. Fetal genetic diagnosis by chorionic villus sampling: evaluation of the five-year experience from a single center. Fetal Pediatr Pathol. 2021;40:281-9.
2. Practice Bulletin No. 162: Prenatal diagnostic testing for genetic disorders. Obstet Gynecol. 2016;127:e108-e22.
3. Ghi T, Sotiriadis A, Calda P, et al; International Society of Ultrasound in Obstetrics and Gynecology (ISUOG). ISUOG Practice Guidelines: invasive procedures for prenatal diagnosis. Ultrasound Obstet Gynecol. 2016;48:256-68.
4. Navaratnam K, Alfirevic Z; Royal College of Obstetricians and Gynaecologists. Amniocentesis and chorionic villus sampling: Green-top Guideline No. 8 July 2021: Green-top Guideline No. 8. BJOG. 2022;129:e1-e15.
5. Torun R, Sever B, Tuncer Can S, Sağlam C, Özer M, Emiralioğlu Çakır Z. Analysis of amniocentesis results in a tertiary care center: a retrospective cohort study. Forbes J Med. 2024;5:129-34.
6. Dikensoy E, Cebesoy FB, Öztürk E, Uğur MG, Balat Ö, Kutlar Aİ. Initial chorionic villus sampling results of Gaziantep University: one year analysis. Turk J Obstet Gynecol. 2010;7:267-72.
7. Jorge P, Mota-Freitas MM, Santos R, Silva ML, Soares G, Fortuna AM. A 26-year experience in chorionic villus sampling prenatal genetic diagnosis. J Clin Med. 2014;3:838-48.
8. Bilen E, Yüksel M, Köse S, Tola E, Sezik M. Gebeliğin 11-14 haftasında karyotipleme amaçlı koryon villüs biyopsisi: 42 olgunun değerlendirilmesi. Süleyman Demirel Üniversitesi Sağlık Bilimleri Dergisi. 2015;6:1-3.
9. Kuyucu M, Erdogan KM, Adiyaman D, et al. Consecutive 5-year outcomes of chorionic villus sampling at a tertiary center. Medicine (Baltimore). 2025;104:e41582.
10. Chevreau J, Becart L, Sergent F, Foulon A, Gondry J, Jedraszak G. Risk factors for failed chorionic villus sampling: results of a 4-year retrospective study. J Matern Fetal Neonatal Med. 2022;35:141-6.
11. Akolekar R, Beta J, Picciarelli G, Ogilvie C, D'Antonio F. Procedure-related risk of miscarriage following amniocentesis and chorionic villus sampling: a systematic review and meta-analysis. Ultrasound Obstet Gynecol. 2015;45:16-26.
12. Martins AT, Francisco C, Correia H, Cohen Á. Chorionic villus sampling: 10 years of experience in a university referral center. J Gynecol Obstet Hum Reprod. 2020;49:101715.
13. Mujezinovic F, Alfirevic Z. Procedure-related complications of amniocentesis and chorionic villous sampling: a systematic review. Obstet Gynecol. 2007;110:687-94. Erratum in: Obstet Gynecol. 2008;111:779.
14. AIUM-ACR-ACOG-SMFM-SRU practice parameter for the performance of standard diagnostic obstetric ultrasound examinations. J Ultrasound Med. 2018;37:E13-E24.
15. International Society of Ultrasound in Obstetrics and Gynecology; Bilardo CM, Chaoui R, Hyett JA, et al. ISUOG Practice Guidelines (updated): performance of 11-14-week ultrasound scan. Ultrasound Obstet Gynecol. 2023;61:127-43.
16. Sikkema-Raddatz B, Bouman K, Verschuuren-Bemelmans CC, et al. Four years' cytogenetic experience with the culture of chorionic villi. Prenat Diagn. 2000;20:950-5.
17. Manegold-Brauer G, Kang Bellin A, Hahn S, et al. A new era in prenatal care: non-invasive prenatal testing in Switzerland. Swiss Med Wkly. 2014;144:w13915.
18. Manegold-Brauer G, Berg C, Flöck A, Rüland A, Gembruch U, Geipel A. Uptake of non-invasive prenatal testing (NIPT) and impact on invasive procedures in a tertiary referral center. Arch Gynecol Obstet. 2015;292:543-8.
19. Dungan JS, Klugman S, Darilek S, et al. Noninvasive prenatal screening (NIPS) for fetal chromosome abnormalities in a general-risk population: an evidence-based clinical guideline of the American College of Medical Genetics and Genomics (ACMG). Genet Med. 2023;25:100336. Erratum in: Genet Med. 2023;25:100874.
20. Tuncer Can S, Golbasi H, Bayraktar B, et al. Prenatal findings and perinatal outcomes of rare autosomal trisomies: a retrospective cohort study. Fetal Pediatr Pathol. 2025;44:357-71.

21. Benn P, Malvestiti F, Grimi B, Maggi F, Simoni G, Grati FR. Rare autosomal trisomies: comparison of detection through cell-free DNA analysis and direct chromosome preparation of chorionic villus samples. Ultrasound Obstet Gynecol. 2019;54:458-67.
22. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin No. 88, December 2007. Invasive prenatal testing for aneuploidy. Obstet Gynecol. 2007;110:1459-67.
23. Beksaç MS, Unal C, Tanacan A, Fadiloglu E, Çakar AN. Chorionic villus sampling experience of a reference perinatal medicine center. Ann Hum Genet. 2020;84:229-34.
24. Hopkins MK, Dugoff L, Kuller JA. Spinal muscular atrophy: inheritance, screening, and counseling for the obstetric provider. Obstet Gynecol Surv. 2021;76:166-9.
25. Turkey SMA Foundation. For combating SMA disease and healthy tomorrows. Available from: https://www.sma.org.tr/en/carrier-scan
26. Sun Y, Kong X, Zhao Z, Zhao X. Mutation analysis of 419 family and prenatal diagnosis of 339 cases of spinal muscular atrophy in China. BMC Med Genet. 2020;21:133.
27. Shariati M, Davoudi A, Boostani R, et al. The prevalence of SMN gene deletion/duplication in spinal muscular atrophy families referred to neuro-genetic centers of Mashhad, Iran. Egypt J Med Hum Genet. 2024;25:87.
28. Mežnarić S, Belančić A, Rački V, Vitezić D, Mršić-Pelčić J, Pilipović K. Prenatal management of spinal muscular atrophy in the era of genetic screening and emerging opportunities in ın utero therapy. Biomedicines. 2025;13:1796.
29. Beksac MS, Yurter H, Tanacan A, et al. Gestational outcomes of pregnant women who have had invasive prenatal testing for the prenatal diagnosis of spinal muscular atrophy. S Afr J Obstet Gynaecol 2018;24:15.
30. Papp C, Beke A, Mezei G, Tóth-Pál E, Papp Z. Chorionic villus sampling: a 15-year experience. Fetal Diagn Ther. 2002;17:218-27.
31. Akolekar R, Bower S, Flack N, Bilardo CM, Nicolaides KH. Prediction of miscarriage and stillbirth at 11-13 weeks and the contribution of chorionic villus sampling. Prenat Diagn. 2011;31:38-45.
32. Zelig CM, Knutzen DM, Ennen CS, Dolinsky BM, Napolitano PG. Chorionic villus sampling, early amniocentesis, and termination of pregnancy without diagnostic testing: comparison of fetal risk following positive non-invasive prenatal testing. J Obstet Gynaecol Can. 2016;38:441-5.e2.
33. Movahedi M, Farahbod F, Zarean E, Hajihashemi M, Haghollahi F, Farahmand M. Evaluation of fetal and maternal outcomes in chorion villus sampling (CVS). Adv Biomed Res. 2023;12:133.
34. Shahbazian N, Barati M, Arian P, Saadati N. Comparison of complications of chorionic villus sampling and amniocentesis. Int J Fertil Steril. 2012;5:241-4.
35. Vera-Ponce VJ, Loayza-Castro JA, Ballena-Caicedo J, Valladolid-Sandoval LAM, Zuzunaga-Montoya FE, Gutierrez De Carrillo CI. Global prevalence of preeclampsia, eclampsia, and HELLP syndrome: a systematic review and meta-analysis. Front Reprod Health. 2025;7:1706009.
36. Kim N, Joo EH, Kim S, et al. Comparative analysis of obstetric, perinatal, and neurodevelopmental outcomes following chorionic villus sampling and amniocentesis. Front Med (Lausanne). 2024;11:1407710.

DOI: 10.4274/forbes.galenos.2026.23540
Forbes J Med 2026;7:33-8

Original Article

Suprapatellar Entry Intramedullary Nailing in Adult Tibial Shaft Fractures: A Retrospective Clinical Study

Ali Murat BAŞAK[1], Ömer Levent KARADAMAR[2], Mehmet TÜRKER[3]

[1]University of Health Sciences Türkiye, Gülhane Training and Research Hospital, Clinic of Orthopedics and Traumatology, Ankara, Türkiye
[2]Döşemealtı State Hospital, Clinic of Orthopedics and Traumatology, Antalya, Türkiye
[3]Private Beyhekim Hospital, Clinic of Orthopedics and Traumatology, Sakarya, Türkiye

Cite as: Başak AM, Karadamar ÖL, Türker M. Suprapatellar entry intramedullary nailing in adult tibial shaft fractures: a retrospective clinical study. Forbes J Med. 2026;7:33-8

ABSTRACT

Objective: The tibia is one of the most commonly fractured long bones and is frequently exposed to high-energy trauma. Selecting an appropriate treatment method is essential for successful outcomes and improved patient quality of life. In recent years, the suprapatellar entry intramedullary nailing technique has gained attention as a promising surgical approach. The aim of this study was to evaluate the clinical and radiological outcomes of suprapatellar intramedullary nailing in the treatment of adult tibial shaft fractures.

Methods: This retrospective study included adult patients with tibial shaft fractures treated with the suprapatellar intramedullary nailing technique. Fractures were classified using the Arbeitsgemeinschaft für Osteosynthesefragen classification system. Both open and closed fractures were included, and open fractures were classified according to the Gustilo–Anderson classification. Clinical and radiological outcomes were evaluated during follow-up.

Results: The mean postoperative fracture healing time was approximately 9 months. Most patients reported minimal anterior knee pain, with only occasional cases of limited knee motion. Observed complications included superficial infections, minor angular deformities, and small limb length discrepancies. Dynamization was required in two patients. The average operative time was approximately 52 minutes.

Conclusion: Suprapatellar intramedullary nailing represents a reliable and effective surgical option for the treatment of adult tibial shaft fractures. The technique may facilitate fracture reduction and result in favorable clinical outcomes with acceptable complication rates.

Keywords: Tibial shaft fracture, intramedullary nailing, suprapatellar entry

Received: 27.12.2025
Accepted: 24.03.2026
Epub: 26.03.2026
Publication Date: 10.04.2026

Corresponding Author:
Ali Murat BAŞAK
University of Health Sciences Türkiye, Gülhane Training and Research Hospital, Clinic of Orthopedics and Traumatology, Ankara, Türkiye
muratalibasak@gmail.com
ORCID ID: 0000-0003-3379-6052

INTRODUCTION

Tibial shaft fractures are among the most common long bone fractures in the adult population and are predominantly the result of high-energy trauma.[1] The primary goals in the treatment of these fractures are to achieve stable union while preserving limb length, alignment, and rotation, and to restore the patient's functional capacity as early as possible.[2] Accurate classification is essential for treatment planning, and the AO classification system remains the most widely used method in clinical practice (Figure 1).[3]

Traditionally, tibial intramedullary nailing has been performed through an infrapatellar entry portal with the knee in a flexed position. However, the infrapatellar approach has been reported to complicate fracture reduction, particularly in proximal and distal tibial shaft fractures, and to increase the risk of malalignment.[4,5] In addition, this approach has been associated with the development of postoperative anterior knee pain.[5] For these reasons, the suprapatellar entry technique performed with the knee in a semi-extended position has gained increasing attention in recent years.[4,6]

The suprapatellar approach has been reported to offer several surgical advantages, including easier fracture reduction, improved fluoroscopic visualization, and better

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

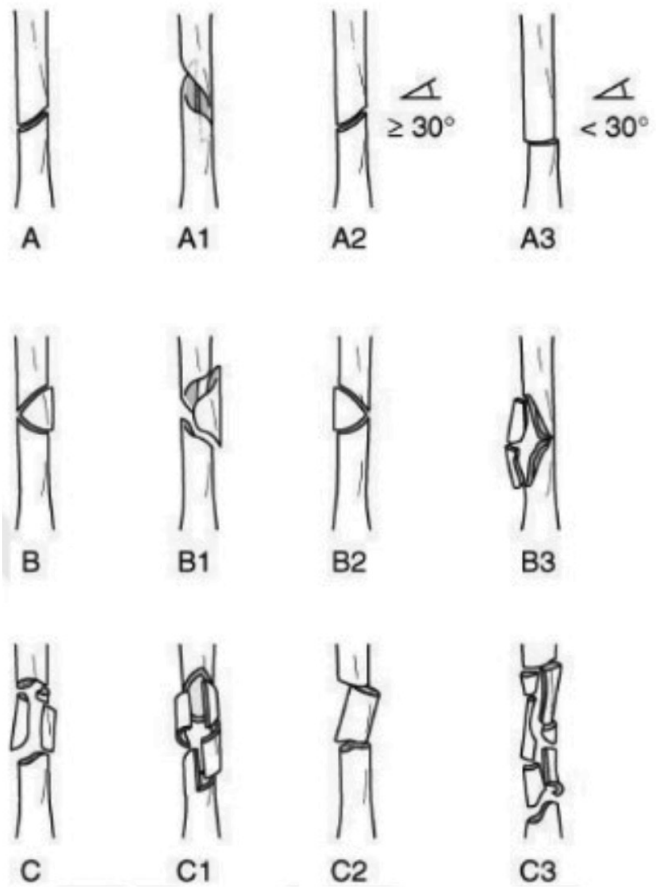


Figure 1. AO classification of tibial fractures

AO: Arbeitsgemeinschaft für Osteosynthesefragen

control of the mechanical axis.[6,7] Clinical studies have demonstrated that suprapatellar nailing may reduce the rate of malalignment compared with the infrapatellar technique while providing at least equivalent functional outcomes.[8,9] Furthermore, recent systematic reviews and meta-analyses support the suprapatellar approach as a safe and effective alternative, particularly for proximal and distal tibial shaft fractures, demonstrating lower malalignment rates and comparable clinical outcomes.[9,10]

The aim of this study is to evaluate the clinical and radiological outcomes of suprapatellar-entry locked intramedullary nailing for the treatment of adult tibial shaft fractures and to discuss the findings in the context of the current literature. Particular emphasis was placed on coronal and sagittal alignment parameters, functional outcomes, and the incidence of anterior knee pain, which remain critical concerns in intramedullary nailing and are frequently addressed in comparison with the traditional infrapatellar approach.

METHODS

Study Design and Patient Selection

This retrospective observational study was conducted at a single tertiary referral center. A total of 25 adult patients with tibial shaft fractures who underwent suprapatellar intramedullary were including in the study. Patients younger than 18 years, those with fractures exhibiting intra-articular extension, those with type III open fractures, or those with pseudoarthrosis following surgery were excluded from the study. All eligible patients were invited for follow-up evaluation, and the 25 patients who attended the final assessment constituted the study cohort. Among the patients, 7 were female and 18 were male. The mean age was 45 years (range: 18–68) among female patients and 40.1 years (range: 20–75) among male patients. Two fractures were open, and 23 were closed. According to the Gustilo–Anderson classification, one open fracture was type I and another was type II. Fractures were classified according to the AO classification system as type A in 14 patients (56%), type B in 6 patients (24%), and type C in 5 patients (20%). Based on fracture location, 7 fractures (28%) were proximal, 10 (40%) were mid-shaft, and 8 (32%) were distal. Radiological evaluation included assessment of coronal and sagittal alignment on standard anteroposterior and lateral radiographs. Malalignment was defined as an angular deformity of ≥5° in any plane. Rotational alignment was assessed clinically and radiographically by comparing limb positioning and cortical continuity. Leg length discrepancy was measured using orthoradiographic imaging, and shortening of ≥1 cm was recorded. All radiographic evaluations were performed at final follow-up. With regard to associated injuries, one patient had a rib fracture, another had a distal radius fracture, and a third had a proximal humerus fracture. The remaining patients had isolated tibial fractures. The patient with a rib fracture did not have a pulmonary contusion; therefore, reamed intramedullary nailing was not contraindicated. All patients underwent surgery within the first week following injury.

The study protocol was approved by the Sakarya University Ethics Committee (approval number: 71522473/050.01.04/35, date: 28.02.2017).

Radiological Evaluation

Radiological assessment was performed using standardized anteroposterior and lateral radiographs at the final follow-up. Coronal and sagittal alignment was evaluated by measuring the angulation between proximal and distal fracture fragments.

Malalignment was defined as an angular deformity of ≥5° in either the coronal or sagittal plane. Rotational alignment

was assessed clinically by comparing the foot progression angle and limb symmetry, and radiographically by evaluating cortical continuity.

Leg length discrepancy was measured using orthoradiographic imaging, and shortening of ≥1 cm was considered clinically relevant.

Surgical Procedure

All procedures were performed under general, spinal, or epidural anesthesia without the application of a tourniquet. With the knee in approximately 10–15° of flexion (semi-extended position), a longitudinal incision approximately 3 cm in length was made proximal to the patella (Figure 2). Access to the patellofemoral joint was achieved through the quadriceps tendon, and a cannula was introduced into the anterior cortex of the proximal tibia under fluoroscopic guidance. A guidewire was advanced across the fracture site to the distal tibia under fluoroscopic control (Figure 3). After appropriate positioning was confirmed, sequential reaming of the medullary canal was performed, and a preselected locked intramedullary nail was inserted. Reduction and alignment were assessed fluoroscopically. Distal locking screws were placed using the free-hand technique, followed by placement of proximal locking screws. All fractures were treated with closed reduction and statically- or dynamically- locked nails as indicated.

Postoperative Management and Follow-up

Postoperatively, no splinting was applied. Weight-bearing was gradually increased based on fracture characteristics and patient tolerance. Full weight- bearing without support was allowed after radiological evidence of union. The mean postoperative hospital stay was 4 days (range: 2–8 days). Patients were followed monthly with clinical and radiographic evaluations.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows software. Continuous variables were expressed as mean ± standard deviation, and categorical variables as number and percentage. The normality of data distribution was assessed using the Shapiro–Wilk test. Comparisons between male and female patients regarding fracture healing time, follow-up duration, Lysholm score, and Cincinnati score were performed using the independent samples t-test. A p value <0.05 was considered statistically significant. Parametric tests were applied because the continuous variables were normally distributed according to the Shapiro–Wilk test.

RESULTS

The clinical and radiological outcomes of the 25 patients who attended the final follow-up evaluation were analyzed. The average follow-up duration of our patients was 14.8 months (range: 6 to 24 months). The average fracture union time was 9.14 months (range: 6 to 15 months) (Table 1).

The patients were evaluated according to the Lysholm and Cincinnati functional scoring systems. The average Lysholm score was 90.8 (range: 64-100), and the average Cincinnati score was 26 (range: 15-30). Four patients experienced minimal patellofemoral pain, while the remaining patients experienced none. The patients were also assessed for joint range of motion limitations. Four patients had 110 degrees of knee flexion. Based on these results, all patients reported being satisfied with the surgery (Table 2).

Table 1. Some descriptive data of the patients

	Min	Max	Mean ± SD
Age (year)	18	75	42.6±16.33
Healing time (month)	6	15	9.14±2.98
Follow-up period (Month)	6	24	14.83±5.30
Lysholm	64	100	90.88±9.30
Cincinnati	15	30	26.83±3.81

SD: Standard deviation, Min: Minimum, Max: Maximum

Table 2. Comparison of healing time, follow-up duration, Lysholm, and Cincinnati scores according to gender

	Male	Female	p value
Healing time (month)	9.19±3.29	9±2	0.90
Follow-up period (Month)	15.18±4.29	14±7.59	0.63
Lysholm	91.06±10.59	90.43±5.68	0.88
Cincinnati	27.71±3.01	24.7±5.09	0.08



Figure 2. (a) Marking of the incision site. (b) Making the suprapatellar incision. (c) Surgical painting and draping procedures

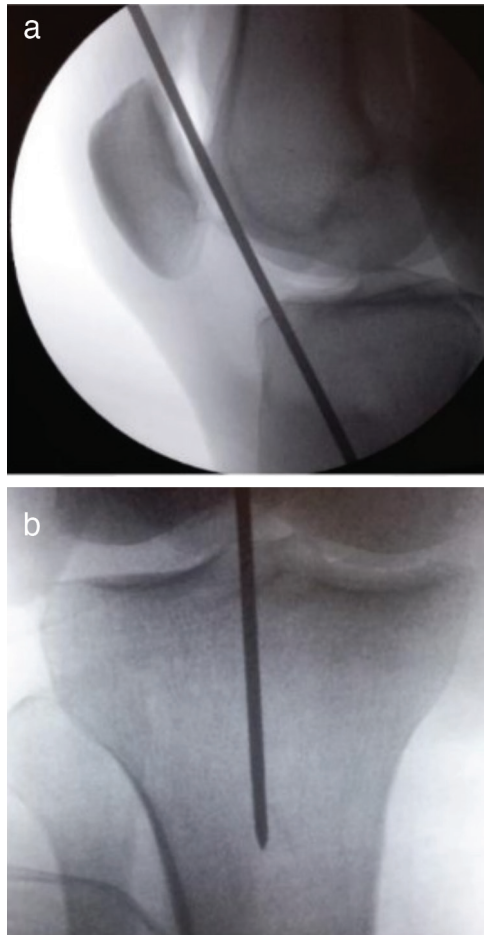


Figure 3. (a) The lateral fluoroscopic image of the guidewire sent as suprapatellar. (b) The anterior posterior fluoroscopic image of the guidewire sent as suprapatellar

Two patients reported discomfort only while sitting with their knees flexed. None of these patients developed compartment syndrome. In 1 patient (4%), a valgus deformity with an angulation between 5° and 10° was observed. Anterior displacement of the distal fragment of over 5° occurred in only 1 patient (4%); however, the patient had no clinical complaints. Only 1 patient (4%) had a 1 cm leg-length discrepancy, which was detected incidentally on an orthoradiogram. The patient had no clinical complaints. The functional outcome scores of patients with radiological malalignment or leg-length discrepancy were analyzed individually. These patients had Lysholm and Cincinnati scores within the overall cohort range and did not exhibit a statistically significant deterioration in functional outcome. The minor radiological deviations detected did not translate into clinically relevant functional limitations at the final follow-up. Superficial infections developed in only 2 cases (8%). These were open fractures, and the infections resolved with antibiotic therapy.

DISCUSSION

The principal finding of this study is that suprapatellar intramedullary nailing provided satisfactory coronal and sagittal alignment, favorable functional outcomes, and a low incidence of anterior knee pain in adult tibial shaft fractures.

Comparisons of suprapatellar and infrapatellar approaches for tibial intramedullary nailing have demonstrated broadly similar clinical outcomes, with differences primarily in alignment control and knee-related symptoms.[10] Meta-analytical data further suggest that the suprapatellar approach may facilitate maintenance of fracture reduction, particularly in fractures with a tendency toward malalignment.[11] In our study, satisfactory radiological alignment was achieved in the majority of patients, supporting the effectiveness of the suprapatellar technique in maintaining fracture reduction.

Functional outcomes following tibial shaft fracture fixation have been extensively evaluated in the literature. Systematic reviews have shown no significant differences in patient-reported outcome measures between suprapatellar and infrapatellar approaches.[12,13] Consistent with these findings, patients in our series achieved acceptable functional recovery without clinically relevant limitations in activities of daily living.

Anterior knee pain remains one of the most commonly reported complaints after tibial intramedullary nailing. Several comparative studies have suggested that suprapatellar entry may be associated with a lower incidence of anterior knee pain than infrapatellar entry.[14] Clinical investigations have also indicated that preservation of the extensor mechanism and avoidance of patellar tendon violation may reduce postoperative knee discomfort.[15] In our cohort, anterior knee pain was infrequent and mild, which further supports the potential benefit of the suprapatellar entry technique.

Rotational alignment is a critical determinant of long-term functional outcome after tibial shaft fracture fixation. Recent studies have demonstrated a lower prevalence of rotational malalignment with suprapatellar nailing compared with infrapatellar techniques.[16] Improved limb control and fluoroscopic visualization in the semi-extended position are thought to contribute to this advantage. In our study, no clinically significant rotational deformity was detected, consistent with these observations.

Concerns regarding the risk of knee sepsis associated with intra-articular instrumentation during suprapatellar nailing have been addressed in several clinical studies. Available evidence indicates that suprapatellar nailing is not associated with an increased incidence of postoperative knee sepsis, even in open tibial fractures.[17,18] In our study, no cases of septic arthritis occurred; superficial infections were limited in number and resolved with appropriate treatment.

Recent systematic reviews comparing suprapatellar, parapatellar, and infrapatellar techniques have concluded that suprapatellar nailing is a safe and effective option with acceptable complication rates and reliable alignment control.[19,20] Furthermore, contemporary reviews and technical descriptions have emphasized that refinements in surgical technique and instrumentation have enhanced the safety and reproducibility of the suprapatellar approach.[21] In our experience, the use of modern suprapatellar instrumentation facilitated accurate nail insertion and stable fixation, which likely contributed to the favorable clinical and radiological outcomes .

This study has several methodological limitations that should be acknowledged. First, the retrospective design may introduce potential selection bias and limit the ability to establish causal relationships. Second, the relatively small sample size may reduce statistical power and restrict the generalizability of the findings. Third, the absence of a comparative control group precludes direct comparison of the suprapatellar approach with alternative techniques, particularly the infrapatellar approach.

CONCLUSION

Suprapatellar entry intramedullary nailing provided satisfactory coronal and sagittal alignment, favorable functional outcomes, and a low incidence of anterior knee pain in adult tibial shaft fractures. The semiextended position may facilitate fracture reduction and improve intraoperative alignment control. Radiological deviations observed in this series were minimal and did not result in clinically significant functional impairment. Complication rates were acceptable, and no major implant-related or intraarticular infectious complications were observed.

Although the retrospective design and limited sample size restrict the strength of definitive conclusions, the findings of this study suggest that the suprapatellar approach is a reliable and effective surgical option for appropriately selected adult tibial shaft fractures.

Ethics

Ethics Committee Approval: The study protocol was approved by the Sakarya University Ethics Committee (approval number: 71522473/050.01.04/35, date: 28.02.2017).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.M.B., M.T., Concept: A.M.B., M.T., Design: A.M.B., Ö.L.K., Data Collection or Processing: A.M.B., Analysis or Interpretation: A.M.B., M.T., Literature Search: A.M.B., Ö.L.K., Writing: A.M.B., Ö.L.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Larsen P, Elsoe R, Hansen SH, Graven-Nielsen T, Laessoe U, Rasmussen S. Incidence and epidemiology of tibial shaft fractures. Injury. 2015;46:746-50.
2. Duan X, Al-Qwbani M, Zeng Y, Zhang W, Xiang Z. Intramedullary nailing for tibial shaft fractures in adults. Cochrane Database Syst Rev. 2012;1:CD008241.
3. Rudge W, Newman K, Trompeter A. Fractures of the tibial shaft in adults. Orthop Trauma. 2014;28:243-55.
4. Morandi M, Banka T, Gaiarsa GP, et al. Intramedullary nailing of tibial fractures: review of surgical techniques and description of a percutaneous lateral suprapatellar approach. Orthopedics. 2010;33:172-9.
5. Yang CY, Tay ST, Kuo LT. Suprapatellar vs infrapatellar approaches for intramedullary nailing of distal tibial fractures: a systematic review and meta-analysis. J Orthop Traumatol. 2023;24:14.
6. Cinats DJ, Viskontas D, Stone T, et al; Canadian Orthopaedic Trauma Society (COTS) ınvestigators. Infrapatellar versus suprapatellar nailing for fractures of the tibia (INSURT study): a multicentered randomized controlled trial. J Orthop Trauma. 2025;39:537-41.
7. Chan DS, Serrano-Riera R, Griffing R, et al. Suprapatellar Versus infrapatellar tibial nail insertion: a prospective randomized control pilot study. J Orthop Trauma. 2016;30:130-4.
8. MacDonald DRW, Caba-Doussoux P, Carnegie CA, et al. Tibial nailing using a suprapatellar rather than an infrapatellar approach significantly reduces anterior knee pain postoperatively: a multicentre clinical trial. Bone Joint J. 2019;101:1138-43.
9. Metcalf KB, Du JY, Lapite IO, et al. Comparison of infrapatellar and suprapatellar approaches for intramedullary nail fixation of tibia fractures. J Orthop Trauma. 2021;35:e45-e50.
10. Sun Q, Nie X, Gong J, et al. The outcome comparison of the suprapatellar approach and infrapatellar approach for tibia intramedullary nailing. Int Orthop. 2016;40:2611-7.
11. Wang C, Chen E, Ye C, Pan Z. Suprapatellar versus infrapatellar approach for tibia intramedullary nailing: a meta-analysis. Int J Surg. 2018;51:133-9.
12. Sepehri A, You D, Lobo AA, Schneider P, Lefaivre KA, Guy P. Comparison of patient-reported outcomes after suprapatellar versus infrapatellar nailing techniques for tibial shaft fractures: a systematic review and meta-analysis. J Orthop Trauma. 2022;36:e208-e14.
13. Wang Z, Xiong X, Lu Z, Gao Y. A systematic review and meta-analysis comparing suprapatellar versus infrapatellar approach intramedullary nailing for tibal shaft fractures. Eur J Trauma Emerg Surg. 2024;50:383-94.
14. Chen J, Wu L, Zhao H, Xue D. Comparison of suprapatellar versus infrapatellar approach for intramedullary nailing of the tibia shaft fractures: a systematic review and meta-analysis. Asian J Surg. 2023;46:5270-1.
15. Soraganvi PC, Anand-Kumar BS, Rajagopalakrishnan R, Praveen-Kumar BA. Anterior knee pain after tibial intra-medullary nailing: is it predictable? Malays Orthop J. 2016;10:16-20.
16. Alderlieste DS, Cain ME, van der Gaast N, et al. Prevalence of rotational malalignment after infrapatellar versus suprapatellar intramedullary nailing of tibial shaft fractures. JB JS Open Access. 2024;9:e23.00134.
17. Mitchell PM, Weisenthal BM, Collinge CA. No incidence of postoperative knee sepsis with suprapatellar nailing of open tibia fractures. J Orthop Trauma. 2017;31:85-9.
18. Marecek GS, Nicholson LT, Broghammer FH, et al. Risk of knee sepsis after treatment of open tibia fractures: a multicenter comparison of suprapatellar and infrapatellar approaches. J Orthop Trauma. 2018;32:88-92.
19. Lu K, Wu ZQ, Qian RX, Gao YJ. The efficacy of suprapatellar, parapatellar and infrapatellar intramedullary nailing in the treatment of tibial fractures: a systematic review and meta-analysis. Arch Orthop Trauma Surg. 2024;144:4917-27.
20. Ringenberg JD, Tobey JL, Horinek JL, Teague DC. Suprapatellar versus infrapatellar approach for intramedullary nail fixation of tibial shaft fractures: a review of the literature. OTA Int. 2022;5:e196.
21. Collinge CA, Rickert MM, Mitchell PM, Boyce RH. Refined techniques in tibial nailing. J Am Acad Orthop Surg. 2025;33:e291-e300.

DOI: 10.4274/forbes.galenos.2026.33603
Forbes J Med 2026;7:39-45

Original Article

Evaluation of Triglyceride Glucose Index and Systemic Inflammatory Biomarkers in Non-Diabetic Patients with Adhesive Capsulitis: A Cross-Sectional Clinical Study

Dilara EKİCİ ZİNCİRCİ[1], Zeynep ERDEM EL[1], Sevgi ATAR[1], Esma DEMİRHAN[1], Ömer KURU[1], Mehmet ZİNCİRCİ[2]

[1]University of Health Sciences Türkiye, Prof. Dr. Cemil Taşcıoğlu City Hospital, Clinic of Physical Medicine and Rehabilitation, İstanbul, Türkiye
[2]İstanbul University Faculty of Medicine, Department of Algology, İstanbul, Türkiye

Cite as: Ekici Zincirci D, Erdem El Z, Atar S, Demirhan E, Kuru Ö, Zincirci M. Evaluation of triglyceride glucose index and systemic inflammatory biomarkers in non-diabetic patients with adhesive capsulitis: a cross-sectional clinical study. Forbes J Med. 2026;7:39-45

ABSTRACT

Objective: Adhesive capsulitis (AC) is a painful shoulder condition characterized by progressive restriction of active and passive range of motion. This study aimed to compare the triglyceride–glucose (TyG) index and haematology-derived inflammatory indices between patients with AC and healthy controls and to evaluate their associations with the presence of AC.

Methods: In this study, the demographic data, clinical findings and laboratory results of non-diabetic patients with AC were evaluated. Fasting plasma glucose (FPG), lipid profile [triglycerides, high-density lipoprotein (HDL), low-density lipoprotein (LDL), and total cholesterol], C-reactive protein (CRP), vitamins D and B12, and complete blood count, obtained within the preceding 3 months, were recorded. The TyG index, neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and systemic immune-inflammation index (SII) were calculated.

Results: Seventy-one participants were included (38 controls and 33 AC). The AC group was older and had higher body mass index (both p≤0.001). FPG (p=0.010), LDL (p=0.001), total cholesterol (p=0.049), and TyG index (p=0.020) were higher in AC, whereas NLR, PLR, SII, CRP, HDL, vitamin D, and B12 were comparable between groups (all p>0.05). In adjusted models, age was a consistent predictor of AC [adjusted odds ratio (aOR) 1.094, 95% confidence interval (CI) 1.036–1.155; p=0.001], and LDL was independently associated with AC (aOR 1.019, 95% CI 1.001–1.037; p=0.044).

Conclusion: Non-diabetic AC was associated with metabolic dysregulation, while haematology-derived inflammatory indices remained comparable to controls. Beyond symptomatic management, clinicians should consider screening patients with AC for metabolic dysregulation, as this may aid risk stratification and care.

Keywords: Adhesive capsulitis, blood glucose level, bloog triglycerides level, inflammation

Received: 07.12.2025
Accepted: 24.03.2026
Publication Date: 10.04.2026

Corresponding Author:
Dilara EKİCİ ZİNCİRCİ
University of Health Sciences Türkiye, Prof. Dr. Cemil Taşcıoğlu City Hospital, Clinic of Physical Medicine and Rehabilitation, İstanbul, Türkiye
drdilaraekici@gmail.com
ORCID ID: 0000-0001-7702-0227

INTRODUCTION

Adhesive capsulitis (AC) is a painful shoulder disorder characterised by restricted movement.[1] It affects approximately 2–5% of the general population and occurs more frequently among individuals with metabolic and endocrine comorbidities.[2] While AC has traditionally been viewed as a local mechanical and inflammatory disorder, growing evidence suggests that chronic low-grade inflammation and metabolic dysfunction may also contribute to its pathogenesis.[3] The risk of AC is particularly increased in individuals with diabetes mellitus (DM).[4] More broadly, metabolic and endocrine abnormalities have been linked to AC; inflammatory and other metabolic profiles may influence disease development and clinical course.[5]

Studies evaluating glucose-related parameters in AC have yielded inconsistent findings, with some reporting associations with symptom severity and others showing no differences in blood glucose compared with controls.[6-10] In contrast, adverse lipid profiles—particularly higher total cholesterol and triglyceride levels—have been associated

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

with AC.[11,12] Thyroid dysfunction may also contribute to pro-inflammatory and fibrotic pathways relevant to AC.[13] Overall, these observations support further investigation of immunometabolic markers in AC.

The triglyceride–glucose (TyG) index is a practical surrogate marker of insulin resistance, calculated from fasting triglyceride and glucose values, and is readily available in clinical practice.[14,15] It has also been used as an indicator of metabolic syndrome.[16] Haematology-derived inflammatory indices, obtained from routine complete blood count parameters, are increasingly used as accessible markers of systemic inflammation. The neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR) are widely used indices,[17] and the systemic immune-inflammation index (SII), combining neutrophil, lymphocyte and platelet parameters, has also been proposed as a useful marker of systemic inflammation.[18]

It is increasingly emphasised that AC may involve systemic inflammation and metabolic disturbances beyond localised pathology.[2,3] Given that insulin resistance has been linked to increased pain and restricted movement in AC,[2] further studies focusing on the immunometabolic profile are warranted. Accordingly, we compared the TyG index and haematology-based inflammatory indices between non-diabetic patients with AC and age- and sex-matched healthy controls and evaluated their associations with the presence of AC. We hypothesised that non-diabetic patients with AC would have a higher TyG index and a more pro-inflammatory haematological profile than controls, and that higher TyG index and inflammatory markers would be independently associated with AC.

METHODS

This was an observational, cross-sectional, case-control study involving prospective recruitment. It compared non-diabetic patients with AC to age- and sex-matched healthy controls. Approval was obtained from the University of Health Sciences Türkiye, Prof. Dr. Cemil Taşcıoğlu Ethics Committee (approval number: 63, date: 24.02.2025), and written informed consent was obtained from all participants. The principles of the Declaration of Helsinki were adhered to throughout the research process.

Participants

Patients who had attended the Physical Medicine and Rehabilitation outpatient clinics for at least three months between March and November 2025, who presented with complaints of shoulder pain and limited movement, who were clinically diagnosed with AC, and who met the inclusion/exclusion criteria were invited to participate in the study.

Inclusion criteria were: age between 18–70 years; shoulder pain lasting at least 3 months with limited shoulder movement; magnetic resonance imaging of the affected shoulder performed within the last three months; complete blood count, C-reactive protein (CRP), fasting plasma glucose (FPG), and fasting lipid panel [triglycerides, high-density lipoprotein (HDL), low-density lipoprotein (LDL), total cholesterol]; and 25(OH)D and vitamin B12 levels measured within the last three months. Exclusion criteria were: a history of shoulder surgery or shoulder trauma; diagnosis of DM or pre-DM; use of triglyceride-lowering medication (statins, fibrates, omega-3, thiazolidinediones); thyroid disease (thyroidectomy, autoimmune thyroiditis, or thyroid dysfunction); pregnancy; acute inflammation; history of malignancy; alcohol use; chronic kidney disease or active infectious disease; viral hepatitis; and liver cirrhosis.

A total of 192 AC patients were evaluated during the study period. Of these, 159 were excluded from the study because they met exclusion criteria or had missing data (shoulder trauma: 15; DM: 83; use of lipid-lowering medication: 20; lack of current laboratory tests: 35; chronic kidney disease: 3; thyroid dysfunction: 17). The control group was recruited sequentially from the same outpatient clinics and consisted of individuals without shoulder-related symptoms who did not meet the exclusion criteria. Demographic, clinical, and laboratory data were recorded for all cases included in the study.

Calculation of Indices

The TyG index was calculated using the following formula:

TyG index = natural log of
((fasting triglycerides [mg/dL] × FPG [mg/dL]) / 2)

The following complete blood count parameters were used for the assessment of systemic inflammation:

NLR = neutrophil count/lymphocyte count

PLR = platelet count/lymphocyte count

SII = platelet count × neutrophil count/lymphocyte count

Statistical Analysis

Descriptive statistics were expressed as the mean ± standard deviation for continuous variables, and as the number of observations (n) for categorical variables. The normality of the data distribution was tested using the Shapiro–Wilk test. Continuous variables that were normally distributed were compared using the independent-samples t-test, whereas non-normally distributed variables were compared using the Mann–Whitney U test. Categorical variables were compared using the chi-squared test or Fisher's exact test.

To evaluate factors associated with the presence of AC, binary logistic regression analyses were conducted, and the results were reported as adjusted odds ratios (aORs) with 95% confidence intervals (CIs). Given the modest sample size, the number of covariates entered into each multivariable model was restricted, and covariates were selected a priori based on clinical relevance to reduce the risk of overfitting. A series of hierarchical models was constructed: Model 0 included age and BMI; Model 1 included LDL; Model 2 included FPG; and Model 3 included the TyG index. To minimise potential multicollinearity, collinearity diagnostics were assessed (e.g., tolerance and variance inflation factors). B ecause the TyG index is derived from FPG and triglycerides, closely related metabolic variables were not entered simultaneously the same model when appropriate. Model fit was evaluated using the Hosmer–Lemeshow goodness-of-fit test. A two-sided p value <0.05 was considered statistically significant. Analyses were performed using IBM SPSS Statistics for Windows, version 23.0 (IBM Corp., Armonk, NY, USA).

RESULTS

A total of 71 participants were included in the study (38 healthy, 33 AC). The mean age and BMI of the patient group were higher than those of the control group (p<0.001 and p=0.001, respectively). The healthy group had a higher level of education, whereas a higher proportion of the patient group were retired or housewives (p=0.001 and p=0.047, respectively). No significant differences in gender or smoking habits were found between the groups (p=0.967 and p=0.908, respectively). As the variables relating to hand dominance and affected side were only present in the patient group, no intergroup comparisons were made for these variables (see Table 1).

Compared with the healthy group, the patient group had higher FPG, LDL, and total cholesterol levels (p=0.010, p=0.001, and p=0.049, respectively). Additionally, the TyG index was significantly higher in patients (p=0.020). However, no significant differences were observed between the groups with respect to neutrophil, lymphocyte, platelet, NLR, PLR, SII, HDL, vitamin D, vitamin B12, and CRP values (all p>0.05). Triglyceride levels tended to be higher in the patient group, although the difference did not reach statistical significance (p=0.071; see Table 2).

Multivariate logistic regression analyses were performed to identify factors that predict the presence of AC. In Model 0, which included age and body mass index (BMI), age was an independent predictor of illness (aOR =1.094; 95% CI 1.036–1.155; p=0.001), whereas BMI was not significant (p=0.111). In Model 1, which included LDL, age remained significant (aOR =1.085; 95% CI 1.025–1.148; p=0.005), and LDL level was independently associated with the presence of AC (aOR =1.019; 95% CI 1.001–1.037; p=0.044). In Model 2, when FPG was added to LDL, FPG did not show an independent association (p=0.705), but LDL retained its significance (p=0.042). In Model 3, when the TyG index was added alongside LDL, the TyG index did not demonstrate an independent association (p=0.299), while the association of LDL declined to borderline significance (aOR =1.017; 95% CI 0.999–1.036; p=0.064). Age remained a consistent independent predictor of the presence of AC in all models (p≤0.013) (Table 3).

DISCUSSION

This study examined whether the TyG index, lipid profile, and haematology-derived inflammatory indices differ between non-diabetic patients with AC and healthy controls, and whether these markers are associated with AC. Overall, our findings indicate that non-diabetic AC have a more adverse metabolic profile, whereas haematology-derived inflammatory indices were comparable to controls. In regression analyses, age remained associated with AC across adjusted models. Although LDL showed a significant association with AC in the age-adjusted analysis, this association was no longer statistically significant after adjustment for the TyG index, indicating that the LDL–AC association may overlap with insulin resistance-related metabolic factors rather than representing a stable independent risk factor.

DM, thyroid disorders, and dyslipidaemia have been reported as risk factors for AC.[19,20] However, because dyslipidaemia commonly coexists with diabetes and thyroid disease, its independent contribution is difficult to isolate.[19,21] Evidence is also mixed: Zhang et al.[21] (including Mendelian randomisation analyses) found no significant association between circulating lipid levels and AC, whereas case-control data suggest that higher lipid levels may be related to AC, although triglycerides may not differ consistently from controls.[7,22] In our cohort, LDL and total cholesterol were higher in patients with AC, whereas triglyceride levels were comparable between groups. Moreover, the LDL–AC association was evident only in age-adjusted analyses and was attenuated after accounting for the TyG index, underscoring the importance of insulin resistance–related metabolic context when interpreting lipid associations. Differences across studies may reflect design- and cohort-related factors, including medication use and residual confounding.

Hyperglycaemia may promote connective tissue changes via glycation and may be linked to heightened oxidative and inflammatory signalling.[23,24] Several studies have associated higher FPG with AC risk,[25,26] although much of this evidence derives from diabetic cohorts. In non-diabetic populations, higher FPG within the normoglycaemic range has also been

Table 1. Comparison of demographic and clinical characteristics between the healthy and patient groups						
		Healthy (38)	**Patient (33)**	**p value**	**ES**	**CI**
Age (years) (mean ± SD)		42.08±12.43	56.73±11.11	<0.001[M]	0.520	0.33–0.67
Sex (female), n		24	21	0.967[χ]	0.005	-0.228–0.238
BMI (kg/m²) (mean ± SD)		25.42±3.42	28.33±4.20	0.001[M]	0.410	0.19–0.59
Education	**Primary and middle school, n**	14	25	0.001[χ]	0.39	0.17–0.57
	High school, n	24	8			
Occupation	**Retired/household, n**	12	18	0.047[χ]	0.29	-
	Office work, n	20	8			
	Manual labor, n	6	7			
Hand dominans (right), n		-	32	-	-	-
Affected side	**Dominant, n**	-	21	-	-	-
	Non-dominant, n	-	12			
Smoking status (current), n		12	10	0.908	0.014	-0.23–0.23
VAS activity (mean ± SD)		-	5.73±2.44	-	-	-
VAS rest (mean ± SD)		-	1.39±1.46	-	-	-
VAS night (mean ± SD)		-	4.91±2.71	-	-	-

Data are presented as mean ± SD for continuous variables and as n for categorical variables [M]Mann–Whitney U test; [χ]chi-square test were used

BMI: Body mass index, VAS: Visual analog scale, n: Number, CI: Confidence interval, ES: Effect size, SD: Standard deviation Variables related to hand dominance and the affected side were available only in the patient group; therefore. no between-group statistical comparisons were performed for these variables (—)

Table 2. Comparison of laboratory parameters between the healthy and patient groups					
	Healthy (38)	**Patient (33)**	**p value**	**ES**	**95% CI**
Neutrophil (Neu) (×10⁹/L)	4.16±1.09	4.39±1.30	0.408[t]	0.196	-0.658–0.267
Lymphocyte (Lym) (×10⁹/L)	2.39±0.80	2.42±0.98	0.845[M]	0.023	-0.211–0.255
Platelets (Plt) (×10⁹/L)	254.84±75.94	278.09±85.88	0.156[M]	0.168	-0.068–0.386
Neu/Lym (NLR)	1.91±0.74	2.16±1.75	0.854[M]	0.022	-0.213–0.254
Plt/Lym (PLR)	117.05±50.08	125.17±42.70	0.203[M]	0.150	-0.370–0.080
SII (SIII)	488.36±245.09	559.18±336.19	0.170[M]	0.163	-0.073–0.382
Fasting plasma glucose (mg/dL)	90.55±11.38	98.82±16.66	0.010[M]	0.305	0.078–0.503
Triglycerides (mg/dL)	110.37±48.60	138.76±64.69	0.071[M]	0.214	-0.020–0.426
HDL (mg/dL)	57.71±17.14	53.45±13.15	0.250[t]	0.273	-0.191–0.736
LDL (mg/dL)	111.61±27.88	146.67±47.90	**0.001**[t]	0.901	-1.383–(-0.413)
Total cholesterol (mg/dL)	188.42±34.71	211.67±57.79	**0.049**[t]	0.491	-0.957–(-0.021)
Vitamin D (µg/L)	19.65±7.44	19.94±9.82	0.822[M]	0.027	-0.208–0.258
Vitamin B12 (ng/L)	346.89±109.31	298.76±138.21	0.109[t]	0.385	-0.085–0.852
TyG index	8.31±0.79	8.72±0.52	0.020[M]	0.276	0.046–0.479
CRP (mg/L)	2.31±2.14	3.32±2.74	0.071[M]	0.214	-0.020–0.426

Data are presented as mean ± SD. p values were obtained using Mann–Whitney U test ([m]) or independent-samples t-test ([t]), as appropriate. ES indicates ES (r for Mann–Whitney U; |Hedges' g| for t-test). 95% CI denotes the 95% CI of the ES

NLR: Neutrophil-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, SII: Systemic immune-inflammation index, LDL: Low-density lipoprotein, HDL: High-density lipoprotein, TyG: Triglyceride–glucose index, CRP: C-reactive protein, SD: Standard deviation, ES: Effect size, CI: Confidence interval

associated with shoulder pathology and AC.[27,28] Consistent with these reports, FPG levels were higher in our non-diabetic AC group than in controls.

The TyG index, a practical surrogate marker of insulin resistance derived from fasting glucose and triglyceride levels, is feasible for routine clinical use and epidemiological research.[14] To our knowledge, this is the first study to assess TyG in patients with AC, and we observed higher TyG values in the AC group than in controls. Limited prior evidence supports a link between insulin resistance and AC severity, with higher homeostasis model assessment of insulin resistance reported to be associated with greater pain, disability, and reduced shoulder mobility.[10] In our regression analyses, the lack of an independent TyG association may reflect limited statistical power and/or shared variance with correlated metabolic measures, underscoring the need for cautious, model-aware interpretation. Overall, these findings suggest that insulin resistance–related metabolic alterations may be relevant in AC even when diabetes is excluded

Inflammation is implicated in AC pathogenesis, with cytokine-mediated processes contributing to fibrosis and abnormal tissue repair.[29-31] Haematology-derived indices such as NLR, PLR and SII are widely used as accessible, non-specific markers of systemic inflammation,[17] but data in AC are limited. Prior studies have reported differences in these indices across AC subtypes or related shoulder conditions.[32,33] In our study, haematology-based inflammatory markers did not differ between patients with AC and controls, suggesting that any inflammatory signal in AC may not be reliably detected by peripheral blood indices. Given the measures' non-specificity and susceptibility to residual variability, limited power may also have contributed to null findings. Future studies incorporating cytokine-based biomarkers and/or tissue-level assessments may better characterise the inflammatory profile of AC.

This study has limitations. First, FPG was based on a single measurement, and repeated assessments would better account for within-person variability. Second, the single-centre design and modest sample size limited the number of covariates that could be included in multivariable models, thereby reducing the precision of estimates. Therefore, we used a streamlined, clinically informed modelling strategy and interpreted results as model-specific. Finally, laboratory values obtained within the preceding three months may have introduced variability due to timing differences relative to clinical evaluation.

Table 3. Factors predicting the presence of adhesive capsulitis: multivariable logistic regression models

Variable	Model 0: age + BMI	Model 1: age + BMI + LDL	Model 2: age + BMI + LDL + FPG	Model 3: age + BMI + LDL + TyG
Age (years) aOR (95% CI) p	1.094 (1.036–1.155) **0.001**	1.085 (1.025–1.148) **0.005**	1.080 (1.016–1.147) **0.013**	1.079 (1.020–1.141) **0.008**
BMI (kg/m²) aOR (95% CI) p	1.129 (0.973–1.310) 0.111	1.101 (0.947–1.281) 0.211	1.100 (0.945–1.280) 0.217	1.091 (0.936–1.271) 0.264
LDL (mg/dL) aOR (95% CI) p	—	1.019 (1.001–1.037) **0.044**	1.019 (1.001–1.037) **0.042**	1.017 (0.999–1.036) 0.064
FPG (mg/dL) aOR (95% CI) p	—	—	1.008 (0.967–1.051) 0.705	—
TyG index aOR (95% CI) p	—	—	—	1.836 (0.583–5.782) 0.299

Dependent variable: patient =1, healthy =0. aOR: Adjusted odds ratio, CI: Confidence interval. BMI: Body mass index, LDL: Low-density lipoprotein, FPG: Fasting plasma glucose, TyG: Triglyceride–glucose index.

CONCLUSION

These findings support consideration of a broader immunometabolic perspective in AC. In clinical practice, evaluating and addressing metabolic risk, particularly insulin resistance–related profiles, may be relevant even in non-diabetic patients and could inform holistic management beyond symptomatic treatment. Further research is needed larger longitudinal studies incorporating direct measures of insulin resistance and more specific inflammatory biomarkers to confirm these associations and clarify the underlying mechanisms.

Ethics

Ethics Committee Approval: Approval was obtained from the Univeristy of Health Sciences Türkiye, Prof. Dr. Cemil Taşcıoğlu Ethics Committee (approval number: 63, date: 24.02.2025).

Informed Consent: Written informed consent was obtained from all participants.

Footnotes

Authorship Contributions

Surgical and Medical Practices: D.E.Z., Z.E.E., S.A., E.D., M.Z., Concept: D.E.Z., S.A., Ö.K., M.Z., Design: D.E.Z., S.A., Ö.K., Data Collection or Processing: D.E.Z., Z.E.E., S.A., E.D., M.Z., Analysis or Interpretation: D.E.Z., Z.E.E., S.A., E.D., Ö.K., M.Z., Literature Search: D.E.Z., Z.E.E., S.A., E.D., M.Z., Writing: D.E.Z., E.D., Ö.K., M.Z.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Millar NL, Meakins A, Struyf F, et al. Frozen shoulder. Nat Rev Dis Primers. 2022;8:59.
2. Hamed-Hamed D, Rodríguez-Pérez C, Pruimboom L, Navarro-Ledesma S. Influence of the metabolic and inflammatory profile in patients with frozen shoulder - systematic review and meta-analysis. BMC Musculoskelet Disord. 2025;26:475. Erratum in: BMC Musculoskelet Disord. 2025;26:664.
3. Navarro-Ledesma S. Frozen shoulder as a systemic immunometabolic disorder: the roles of estrogen, thyroid dysfunction, endothelial health, lifestyle, and clinical implications. J Clin Med. 2025;14:7315.
4. Mahmoud TH, El-Sayed MS, Abdelnaeem WG. Association between frozen shoulder and diabetes mellitus: a narrative review. Deraya International Journal for Medical Sciences and Rehabilitation. 2025;1:16-32.
5. Won KS, Kim DH, Sung DH, et al. Clinical correlation of metabolic parameters on 18F-FDG PET/CT in idiopathic frozen shoulder. Ann Nucl Med. 2017;31:211-7.
6. Safran O, El-Haj M, Leibowitz G, et al. Should patients with frozen shoulder be screened for diabetes mellitus? Orthop J Sports Med. 2017;5:2325967117716450.
7. Sung CM, Jung TS, Park HB. Are serum lipids involved in primary frozen shoulder? A case-control study. J Bone Joint Surg Am. 2014;96:1828-33.
8. Salek AK, Mamun MA, Haque MA, et al. Serum triglyceride level in type 2 diabetes mellitus patients with or without frozen shoulder. Bangladesh Med Res Counc Bull. 2010;36:64-7.
9. Mertens MG, Meeus M, Lluch Girbes E, et al. Differences in biomechanical and metabolic factors between patients with frozen shoulder and asymptomatic individuals. A cross-sectional study. Musculoskelet Sci Pract. 2024;72:102980.
10. Pérez-Montilla JJ, Guzmán-García R, Pruimboom L, Navarro-Ledesma S. Does leptin and insulin levels influence pain and disability in subjects with frozen shoulder? A cross-sectional study. Eur J Pain. 2025;29:e70007.
11. Park HB, Gwark JY, Jung J. What serum lipid abnormalities are associated with adhesive capsulitis accompanied by diabetes? Clin Orthop Relat Res. 2018;476:2231-7.
12. Dyer BP, Burton C, Rathod-Mistry T, Blagojevic-Bucknall M, van der Windt DA. Diabetes as a prognostic factor in frozen shoulder: a systematic review. Arch Rehabil Res Clin Transl. 2021;3:100141.
13. de la Serna D, Navarro-Ledesma S, Alayón F, López E, Pruimboom L. A comprehensive view of frozen shoulder: a mystery syndrome. Front Med (Lausanne). 2021;8:663703.
14. Katar M, Demir O. Triglyceride glucose index: as a glycemic control indicator. Harran Üniversitesi Tıp Fakültesi Dergisi. 2025;22:357-63.
15. Guerrero-Romero F, Simental-Mendía LE, González-Ortiz M, et al. The product of triglycerides and glucose, a simple measure of insulin sensitivity. Comparison with the euglycemic-hyperinsulinemic clamp. J Clin Endocrinol Metab. 2010;95:3347-51.
16. Unger G, Benozzi SF, Perruzza F, Pennacchiotti GL. Triglycerides and glucose index: a useful indicator of insulin resistance. Endocrinol Nutr. 2014;61:533-40.
17. Buonacera A, Stancanelli B, Colaci M, Malatino L. Neutrophil to lymphocyte ratio: an emerging marker of the relationships between the immune system and diseases. Int J Mol Sci. 2022;23:3636.
18. Mangoni AA, Zinellu A. The diagnostic role of the systemic inflammation index in patients with immunological diseases: a systematic review and meta-analysis. Clin Exp Med. 2024;24:27.
19. Sarasua SM, Floyd S, Bridges WC, Pill SG. The epidemiology and etiology of adhesive capsulitis in the U.S. Medicare population. BMC Musculoskelet Disord. 2021;22:828.
20. Milgrom C, Novack V, Weil Y, Jaber S, Radeva-Petrova DR, Finestone A. Risk factors for idiopathic frozen shoulder. Isr Med Assoc J. 2008;10:361-4.
21. Zhang Y, Fan W, Wang Y, Dong T, Liu D, Shao Y. The association between primary frozen shoulder and serum lipids may be overestimated: evidence based on retrospective observational studies and Mendelian randomization. Front Endocrinol (Lausanne). 2025;15:1363018.
22. Sun G, Li Q, Yin Y, Fu W, He K, Pen X. Risk factors and predictive models for frozen shoulder. Sci Rep. 2024;14:15261.

23. Esposito K, Nappo F, Marfella R, et al. Inflammatory cytokine concentrations are acutely increased by hyperglycemia in humans: role of oxidative stress. Circulation. 2002;106:2067-72.
24. Reddy GK. Glucose-mediated in vitro glycation modulates biomechanical integrity of the soft tissues but not hard tissues. J Orthop Res. 2003;21:738-43.
25. Austin DC, Gans I, Park MJ, Carey JL, Kelly JD 4th. The association of metabolic syndrome markers with adhesive capsulitis. J Shoulder Elbow Surg. 2014;23:1043-51.
26. Chan JH, Ho BS, Alvi HM, Saltzman MD, Marra G. The relationship between the incidence of adhesive capsulitis and hemoglobin A1c. J Shoulder Elbow Surg. 2017;26:1834-7.
27. Longo UG, Franceschi F, Ruzzini L, Spiezia F, Maffulli N, Denaro V. Higher fasting plasma glucose levels within the normoglycaemic range and rotator cuff tears. Br J Sports Med. 2009;43:284-7.
28. Park HB, Gwark JY, Kam M, Jung J. Association between fasting glucose levels and adhesive capsulitis in a normoglycemic population: a case-control study. J Shoulder Elbow Surg. 2020;29:2240-7.
29. Lho YM, Ha E, Cho CH, et al. Inflammatory cytokines are overexpressed in the subacromial bursa of frozen shoulder. J Shoulder Elbow Surg. 2013;22:666-72.
30. Kabbabe B, Ramkumar S, Richardson M. Cytogenetic analysis of the pathology of frozen shoulder. Int J Shoulder Surg. 2010;4:75-8.
31. Bunker TD, Reilly J, Baird KS, Hamblen DL. Expression of growth factors, cytokines and matrix metalloproteinases in frozen shoulder. J Bone Joint Surg Br. 2000;82:768-73.
32. Tosun MF, Çiçeklidağ M, Bircan R, et al. The role of inflammatory markers in the differential diagnosis of stiff shoulder disease. Jt Dis Relat Surg. 2025;36:415-9.
33. Çelik İ, Çelik H, Ayanoğlu T, Arıkan E, Kayış SA. Lymphocyte-to-monocyte ratio is a good marker of adhesive capsulitis in rotator cuff tears. Acta Medica Alanya 2023;7;226-31.

DOI: 10.4274/forbes.galenos.2026.82356
Forbes J Med 2026;7:46-52

Original Article

Clinical Differentiation of Painful Ophthalmoplegia: A Comparative Analysis of Diagnostic Groups

Hüseyin Nezih ÖZDEMİR, Elvin HASANLI, Figen GÖKÇAY, Neşe ÇELEBİSOY

Ege University Hospital, Department of Neurology, İzmir, Türkiye

Cite as: Özdemir HN, Hasanlı E, Gökçay F, Çelebisoy N. Clinical differentiation of painful ophthalmoplegia: a comparative analysis of diagnostic groups. Forbes J Med. 2026;7:46-52

ABSTRACT

Objective: Painful ophthalmoplegia (PO) is a clinical syndrome with diverse etiologies, often presenting with overlapping features that complicate diagnosis. Early differentiation is essential to guide appropriate investigations and management. To identify clinical examination findings that differentiate the underlying etiologies of PO and to determine features that may aid in clinical decision-making.

Methods: This retrospective observational study included patients with PO evaluated at a tertiary neuro-ophthalmology center between January 2014 and January 2024. Clinical, demographic, and laboratory data were collected. Patients were classified into diagnostic groups based on etiology. Comparative analyses and binary logistic regression were used to identify clinical predictors that distinguish cavernous sinus-orbital apex (CS-OA) inflammatory syndrome from other causes.

Results: A total of 94 patients were included. CS-OA (28.7%) and carotid-cavernous fistula (CCF) (26.6%) were the most common etiologies. Ptosis was more frequent in CS-OA and was identified as an independent predictor (p=0.031). Chemosis and vision loss were more common in non-CS-OA etiologies and were independently associated with these conditions (p=0.041 and p=0.018, respectively). Chemosis was observed in 88% of CCF cases. Vision loss was most frequent in infectious etiologies, particularly invasive fungal sinusitis. Facial hypoesthesia showed a trend toward an association with CS-OA (p=0.073).

Conclusion: Specific clinical findings, particularly ptosis, chemosis, and vision loss, may help differentiate the underlying etiologies of PO. Careful clinical evaluation can provide valuable guidance for diagnostic decision-making, especially in settings with limited access to advanced investigations.

Keywords: Painful ophthalmoplegia, cavernous sinus, carotid-cavernous fistula, orbital diseases

Received: 16.04.2026
Accepted: 01.06.2026
Epub: 03.06.2026
Publication Date: 09.06.2026

Corresponding Author:
Hüseyin Nezih ÖZDEMİR
Ege University Hospital, Department of Neurology, İzmir, Türkiye
huseyinnezihozdemir@gmail.com
ORCID ID: 0000-0002-0651-7276

INTRODUCTION

Painful ophthalmoplegia (PO) refers to a syndrome of periorbital or hemicranial pain with ipsilateral ocular motor nerve palsies.[1] Although it is uncommon, PO carries clinical importance due to the risk of visual and neurological morbidity.[2] Therefore, accurate and timely differentiation of the underlying etiologies is essential.

Different etiologies of PO may present with similar clinical presentations, complicating the diagnostic process.[3] The spectrum of causes includes vascular, inflammatory, neoplastic, and infectious conditions affecting the cavernous sinus, orbital apex, and orbit.[3] Infectious etiologies include bacterial and fungal infections involving the paranasal sinuses and adjacent structures, as well as systemic infections such as syphilis and mycobacterial disease.[4] Neoplastic causes include both primary and metastatic tumors. Vascular disorders encompass intracavernous carotid artery abnormalities, cavernous sinus thrombosis, and abnormal arteriovenous communications. In addition, infiltrative and systemic inflammatory conditions, including hematologic malignancies, connective tissue diseases, and large-vessel vasculitides, may involve these regions.[4] Tolosa-Hunt syndrome is characterized by idiopathic granulomatous inflammation of the cavernous sinus, superior orbital fissure, or orbital apex.[5]

While detailed imaging and laboratory investigations are required for a definitive diagnosis, clinical findings play a crucial role in guiding clinicians toward a provisional diagnosis and selecting appropriate diagnostic tests. The aim of this study was to identify clinical examination

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

findings that differentiate the etiologies of PO and to identify features that may aid in clinical decision-making.

METHODS

Study Design and Participants

This retrospective observational study was conducted at a tertiary neuro-ophthalmology laboratory. Patients with PO were retrospectively identified through a review of medical records between January 2014 and January 2024. PO was defined as the presence of periorbital or hemicranial pain accompanied by ipsilateral ocular motor nerve involvement.[1]

Patients aged 18 years and older who were evaluated, diagnosed, and treated in the laboratory were included in the study. The diagnostic workup for PO was guided by the differential diagnosis and included laboratory and imaging studies. Laboratory evaluation comprised complete blood count, blood chemistry, inflammatory markers (erythrocyte sedimentation rate and C-reactive protein), thyroid function tests, glycemic evaluation including glycated hemoglobin, and autoantibody testing, including markers for vasculitis and connective tissue diseases when clinically indicated. Serological tests for infectious etiologies were performed when appropriate. Imaging studies included contrast-enhanced cranial and orbital magnetic resonance imaging, with vascular imaging when indicated. Lumbar puncture was performed when clinically indicated. Patients without a complete diagnostic workup or with incomplete clinical data were excluded.

Data were extracted from medical records. Demographic and clinical variables included age, sex, and, when present, comorbidities such as diabetes mellitus, vascular risk factors, autoimmune thyroid disease, and other autoimmune disorders. Neuro-ophthalmological examination findings were recorded. Best-corrected visual acuity was measured in decimal units and converted to the logarithm of the minimum angle of resolution, according to the latest literature.[6] Recorded findings included diplopia, ptosis, proptosis, chemosis, visual loss, limitations of extraocular movements (inward, outward, upward, and downward gaze), and facial hypoesthesia.

Patients were classified into diagnostic groups according to the underlying etiology and predominant anatomical involvement. The diagnostic groups included cavernous sinus-orbital apex inflammatory syndrome (CS-OA), carotid-cavernous fistula (CCF), thyroid eye disease (TED), tumors, idiopathic orbital inflammatory disease (IOID), and infections. The CS-OA group comprised patients with inflammatory conditions involving the cavernous sinus and/or orbital apex, including idiopathic cases. Although these entities differed in etiology, they were grouped together because they shared similar inflammatory involvement of the cavernous sinus and/or orbital apex, with overlapping neuro-ophthalmological features. The tumor group included neoplastic lesions involving the cavernous sinus, orbital apex, or orbit. The infection group included patients with bacterial or fungal infections involving the cavernous sinus, orbital apex, or orbit. Final diagnoses were established based on clinical findings, imaging results, laboratory investigations, and follow-up data when available.

The independent variables in the study were demographic characteristics and findings on clinical examination, including age, sex, comorbidities, diplopia, ptosis, proptosis, chemosis, visual loss, limitations of extraocular movements, and facial hypoesthesia. The dependent variable was the diagnostic group.

Visual loss was defined as a reduction in best-corrected visual acuity compared with the normal expected range for age. Ptosis was defined as the upper eyelid margin positioned below the superior corneal limbus or asymmetrically positioned compared with the fellow eye, as observed on clinical examination. Proptosis was defined as anterior displacement of the globe identified clinically and, when available, supported by orbital imaging. Chemosis was defined as conjunctival edema observed on ophthalmological examination. Extraocular movement restriction was defined as a limitation of ocular excursion in the relevant gaze direction, documented during bedside assessment in the cardinal gaze positions. Fever was defined as a body temperature ≥38 °C at presentation or documented during clinical evaluation.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics (version 26, IBM Corp., Armonk, NY, USA). Continuous variables were assessed for distribution and expressed as median and interquartile range (IQR). Categorical variables were presented as numbers and percentages. Comparisons between diagnostic groups were performed using the Kruskal-Wallis test for continuous variables. For categorical variables, the chi-square test or Fisher's exact test was used as appropriate.

Binary logistic regression was performed to identify clinical examination findings associated with CS-OA, compared with non-CS-OA etiologies. Variables were selected for inclusion in the regression model based on clinical relevance and univariate analysis, with a threshold of $p<0.10$. Multicollinearity was assessed before inclusion in the regression model. The results of the regression analysis

were reported as odds ratios (OR) with 95% confidence intervals (CI). Model coefficients were estimated using maximum likelihood methods. A two-sided p value <0.05 was considered statistically significant.

Ethical Aspect

The study was approved by the Ege University Ethics Committee. The approval number was 25-1.1T/38, and the approval date was 23 January 2025. The study was conducted in accordance with the principles of the Declaration of Helsinki. This study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) guidelines. Informed consent was obtained from each participant.

RESULTS

A total of 111 cases of PO were identified during the study period. After excluding nine patients with incomplete diagnostic workup and eight with incomplete clinical data, 94 patients were included in the study. The median age was 57 years (IQR: 42-65), and 51 patients (54.3%) were female. Table 1 summarizes the demographic features of the study cohort.

Of the 94 patients included in the analysis, 27 (28.7%) had CS-OA; 25 (26.6%) had CCF; 11 (11.7%) had TED; 11 (11.7%) had tumors; 10 (10.6%) had IOID; and 10 (10.6%) had infections. The CS-OA group included 15 cases of idiopathic cavernous sinus inflammation, two cases of cavernous sinus thrombosis, one case of cavernous sinus inflammation due to granulomatosis with polyangiitis, and nine cases of orbital apex inflammation. The tumor group included one malignant epithelial tumor, one chondroid tumor, two hemorrhagic pituitary macroadenomas, three meningiomas, one nasopharyngeal carcinoma, one lymphoma, one multiple myeloma, and one suspected metastatic lesion involving the cavernous sinus and orbit, with clival involvement. In the IOID group, eight patients had orbital myositis and two had orbital pseudotumor. The infection group included nine fungal infections (five cases of mucormycosis and four of aspergillosis) and one case of skull-base infection due to *Streptococcus pneumoniae*.

Table 1. Demographic and clinical characteristics of the study population

Variable	Value
Age, years, median (IQR)	57 (42-65)
Sex, female, n (%)	51 (54.3)
DM, n (%)	29 (30.9)
Vascular risk factors other than DM, n (%)	18 (19.1)
Autoimmune thyroid disease, n (%)	16 (17.1)
Autoimmune disease other than thyroid, n (%)	7 (7.4)
Diplopia, n (%)	78 (83.1)
Ptosis, n (%)	30 (31.9)
Proptosis, n (%)	41 (43.6)
Chemosis, n (%)	30 (31.9)
Vision loss, n (%)	50 (53.2)
Restricted inward gaze, n (%)	13 (13.8)
Restricted outward gaze, n (%)	60 (63.8)
Restricted downward gaze, n (%)	41 (43.6)
Restricted upward gaze, n (%)	17 (18.1)
Facial hypoesthesia, n (%)	27 (28.7)
High fever, n (%)	3 (3.2)

IQR: Interquartile range, DM: Diabetes mellitus.

Comparisons of demographic and clinical characteristics across diagnostic groups are presented in Table 2. Briefly, age differed significantly between groups (p=0.006), with patients in the infection group being older, whereas those with IOID tended to be younger. DM was significantly more frequent in the infection group (p=0.001). Autoimmune thyroid disease was strongly associated with TED (p<0.001).

Among the clinical findings, ptosis was more common in the CS-OA group (p=0.010). Proptosis was more frequently observed in the CCF and TED groups, whereas chemosis was predominantly seen in the CCF group (both p<0.001). Vision loss differed significantly between groups (p<0.001), with higher frequencies observed in the TED and infection groups.

Restricted outward gaze was significantly more frequent in CCF and IOID (p<0.001). Facial hypoesthesia was significantly associated with the CS-OA and tumor groups (p=0.002). No significant differences were observed with respect to sex, vascular risk factors other than diabetes, diplopia, restricted inward, upward, or downward gaze, or the presence of fever (all p>0.05).

To determine which clinical findings differentiate non-CS-OA etiologies from CS-OA, a binary logistic regression analysis was performed. Chemosis (OR =6.54, 95% CI: 1.03-41.53, p=0.041) and vision loss (OR =5.33, 95% CI: 1.32-21.42, p=0.018) were independently associated with non-CS-OA etiologies. In contrast, ptosis was negatively associated with non-CS-OA etiologies (OR =0.21, 95% CI: 0.05-0.86, p=0.031), indicating a higher likelihood of CS-OA in patients presenting with ptosis. The logistic regression model demonstrated acceptable goodness-of-fit according to the Hosmer-Lemeshow test (χ^2=4.481, p=0.811), with a Nagelkerke R^2 value of 0.515.

Other variables, including age, diabetes mellitus, autoimmune thyroid disease, proptosis, and restricted outward gaze, were not independently associated with non-CS-OA etiologies. Facial hypoesthesia showed a trend toward association with non-CS-OA etiologies (OR =0.27, 95% CI: 0.06-1.12, p=0.073), although this did not reach statistical significance (Table 3).

Table 2. Comparison of demographic characteristics and clinical examination findings across diagnostic groups in patients with painful ophthalmoplegia

Variable	CS-OA n=27	CCF n=25	Thyroid eye disease n=11	Tumors n=11	IOID n=10	Infections n=10	p value
Sex, female, n (%)	12 (44.4)	17 (68)	7 (63.6)	4 (36.3)	5 (50)	6 (40)	0.426
Age, years, median (IQR)	44 (35-65)	59 (47.5-66.5)	56 (54-58)	46 (38-62)	39 (26-64)	68 (61-73)	**0.006**
DM, n (%)	6 (22.2)	8 (32)	2 (18.2)	3 (27.3)	1 (10)	9 (90)	**0.001**
Vascular risk factors other than DM, n (%)	7 (25.9)	8 (32)	1 (9.1)	1 (9.1)	0 (0)	1 (10)	0.236
Autoimmune thyroid disease, n (%)	2 (7.4)	4 (16)	8 (72.7)	0 (0)	1 (10)	1 (10)	**<0.001**
Autoimmune disease other than thyroid, n (%)	2 (7.4)	2 (8)	1 (9.1)	1 (9.1)	1 (10)	0 (0)	0.980
Diplopia, n (%)	21 (77.8)	23 (92)	9 (81.8)	10 (90.9)	8 (80)	7 (70)	0.574
Ptosis, n (%)	15 (55.6)	3 (12)	2 (18.2)	3 (27.3)	2 (20)	5 (50)	**0.010**
Proptosis, n (%)	4 (14.8)	20 (80)	9 (81.8)	1 (9.1)	5 (50)	2 (20)	**<0.001**
Chemosis, n (%)	2 (7.4)	22 (88)	0 (0)	0 (0)	4 (40)	2 (20)	**<0.001**
Vision loss, n (%)	10 (37.1)	17 (68)	3 (27.2)	7 (63.6)	4 (40)	9 (90)	**<0.001**
Restricted inward gaze, n (%)	4 (14.8)	2 (8)	2 (18.2)	2 (18.2)	1 (10)	2 (20)	0.886
Restricted outward gaze, n (%)	14 (51.9)	23 (92)	2 (18.2)	7 (63.6)	9 (90)	5 (50)	**<0.001**
Restricted downward gaze, n (%)	10 (37)	9 (36)	6 (54.5)	6 (54.5)	4 (40)	6 (60)	0.660
Restricted upward gaze, n (%)	3 (11.1)	7 (28)	2 (18.2)	2 (18.2)	2 (20)	1 (10)	0.727
Facial hypoesthesia, n (%)	13 (48.1)	6 (24)	0 (0)	6 (54.5)	0 (0)	2 (20)	**0.002**
High fever, n (%)	1 (3.7)	0 (0)	1 (9.1)	0 (0)	0 (0)	1 (10)	0.412

CS-OA: Cavernous sinus-orbital apex, CCF: Carotid-cavernous fistula, DM: Diabetes mellitus, IOID: Idiopathic orbital inflammatory disease, IQR: Interquartile range

Table 3. Binary logistic regression analysis of factors associated with non-cavernous sinus-orbital apex etiologies in patients with painful ophthalmoplegia

Variable	OR	95% CI	p value
Age (years)	1.02	0.97-1.06	0.339
Autoimmune thyroid disease	1.26	0.17-9.29	0.815
DM	0.66	0.13-3.40	0.624
Ptosis	0.21	0.05-0.86	**0.031**
Proptosis	2.11	0.43-10.20	0.352
Chemosis	6.54	1.03-41.53	**0.041**
Vision loss	5.33	1.32-21.42	**0.018**
Restricted outward gaze	1.89	0.47-7.57	0.365
Facial hypoesthesia	0.27	0.06-1.12	0.073

OR: Odds ratio, CI: Confidence interval, DM: Diabetes mellitus

DISCUSSION

This study indicated that specific clinical examination findings may help differentiate the underlying etiologies of PO and provide practical clues for the differential diagnosis. In particular, ptosis was more frequently observed in CS-OA, whereas chemosis and vision loss were more suggestive of alternative etiologies. Facial hypoesthesia also showed a trend toward an association with CS-OA.

Ptosis was most frequent in the CS-OA group, and multivariable analysis demonstrated that it was an independent predictor of CS-OA (p=0.031). Involvement of the oculomotor nerve in the cavernous sinus and orbital apex explains the increased frequency of ptosis in patients with CS-OA.[7,8] The oculomotor nerve innervates the levator palpebrae superioris, and its involvement results in impaired eyelid elevation (ptosis).[8] In our cohort, one patient with granulomatosis with polyangiitis and two patients with cavernous sinus thrombosis were included in the CS-OA group. Cavernous sinus involvement due to granulomatosis with polyangiitis has been reported.[9] Cavernous sinus thrombosis is a rare cause of cavernous sinus involvement.[10] The remaining patients were classified as having idiopathic inflammation of the cavernous sinus, known as Tolosa-Hunt syndrome.[7] Idiopathic orbital apex inflammation reflects a similar inflammatory process confined to the orbital apex and is considered part of the same disease spectrum.[11]

Chemosis refers to edema of the conjunctiva, resulting from impaired venous drainage and increased vascular permeability.[12] CCF is an abnormal arteriovenous communication connecting the carotid arterial system to the cavernous sinus and resulting in elevated pressure in the orbital venous system.[12] This venous hypertension impairs the normal drainage of the conjunctival and episcleral veins, resulting in conjunctival edema and the characteristic appearance of chemosis.[13] In our cohort, chemosis was observed in 88% of patients with CCF and identified as an independent predictor of non-CS-OA etiologies in logistic regression analysis (p=0.041). In addition to its diagnostic value, chemosis may have clinical implications. Direct CCFs, defined as high-flow communications between the intracavernous internal carotid artery and the cavernous sinus, are often characterized by an abrupt onset of chemosis and require early intervention.[13-15]

In our cohort, vision loss was most frequent in patients with PO due to infectious etiologies. Among these, invasive fungal sinusitis, particularly mucormycosis and Aspergillus, was the most common cause. Vision loss was also identified as an independent predictor of non-CS-OA etiologies in logistic regression analysis (p=0.018). This is consistent with the literature, in which these organisms are also reported to be the most frequent causes.[16] Fungal sinusitis can lead to vision loss when the infection invades the optic nerve.[17] On the other hand, CCF was the second most common cause of vision loss in our cohort. Vision loss in CCF is caused by venous hypertension, which impairs orbital venous drainage and increases intraocular pressure, ultimately affecting optic nerve function.[12] In addition, chronic venous congestion may compromise retinal and optic nerve perfusion, leading to ischemic optic neuropathy or retinal vascular dysfunction. Delayed diagnosis and persistent venous hypertension may therefore result in irreversible visual impairment.[12]

Although facial hypoesthesia did not reach statistical significance, it showed a trend toward predicting CS-OA in multivariable analysis (p=0.073). This finding can be explained by the involvement of the trigeminal nerve within the cavernous sinus, where its ophthalmic and maxillary divisions are located.[18] Among non-CS-OA etiologies, facial hypoesthesia was most frequent in the tumor group in our cohort. Tumors invading the cavernous sinus may produce similar sensory deficits by affecting the trigeminal nerve within this region.

Ninety percent of patients with infectious etiologies had DM; however, multivariable analysis did not demonstrate its predictive value for non-CS-OA etiologies. This may be explained by the immunosuppressive effects of DM, which impair host defense mechanisms and increase susceptibility to opportunistic infections, including invasive fungal infections.[19,20] Although DM was not independently associated with non-CS-OA etiologies in the multivariable analysis, the high prevalence of DM among patients with infectious etiologies remains clinically important, likely reflecting the increased susceptibility of diabetic patients to invasive fungal infections. Autoimmune thyroid disease was most frequently observed in patients with TED. The underlying mechanism is autoimmune inflammation directed against orbital fibroblasts, leading to expansion of orbital tissues and involvement of the extraocular muscles.[21] Notably, TED may occur in the absence of overt thyroid dysfunction.[22]

Our findings are generally consistent with previous studies reporting that inflammatory cavernous sinus syndromes commonly present with PO accompanied by cranial nerve involvement, particularly ptosis and sensory deficits.[2] Similarly, chemosis and proptosis are characteristic findings in CCF secondary to orbital venous congestion, whereas vision loss has been reported more frequently in invasive fungal infections and in compressive lesions of the orbital apex.[12,17] Previous studies have also emphasized the association between diabetes mellitus and invasive fungal infections, particularly mucormycosis.[17]

Overall, our findings support the existing literature and highlight specific clinical patterns that may assist in the differential diagnosis of PO in routine clinical practice.

These findings may be particularly helpful in busy and resource-limited clinical settings, where immediate access to advanced imaging and laboratory investigations may be restricted. Recognition of specific clinical patterns may assist clinicians in narrowing the differential diagnosis and prioritizing appropriate diagnostic tests. Ultimately, such an approach may facilitate a more efficient diagnostic process through the integration of clinical findings with targeted investigations. Although the regression model demonstrated acceptable goodness-of-fit, the findings should be considered exploratory and interpreted with caution because of the relatively small sample size and subgroup heterogeneity.

Study Limitations

This study has several limitations. First, its retrospective design may have introduced selection and information bias. Second, the study was conducted at a single center, which may limit the generalizability of the findings. In addition, the sample sizes of certain diagnostic subgroups were relatively small, potentially affecting statistical power; therefore, the findings should be interpreted with caution. Finally, the heterogeneity of underlying etiologies may have contributed to variability in clinical presentation.

CONCLUSION

Specific clinical examination findings may assist in differentiating the underlying etiologies of PO. Ptosis may suggest CS-OA inflammatory syndrome, whereas chemosis and vision loss may indicate alternative etiologies. Although definitive diagnosis relies on imaging and laboratory investigations, careful clinical evaluation may help guide the diagnostic process and support more targeted investigations.

Ethics

Ethics Committee Approval: The study was approved by the Ege University Ethics Committee. The approval number was 25-1.1T/38, and the approval date was 23 January 2025.

Informed Consent: Informed consent was obtained from each participant.

Footnotes

Authorship Contributions

Surgical and Medical Practices: H.N.Ö., F.G., N.Ç., Concept: F.G., N.Ç., Design: H.N.Ö., E.H., Data Collection or Processing: E.H., Analysis or Interpretation: H.N.Ö., Literature Search: H.N.Ö., E.H., Writing: H.N.Ö., F.G., N.Ç.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Ertilav E, Akyol A. Evaluation of patients with painful ophthalmoplegia for benign and secondary etiologies. Neuroophthalmology. 2024;48:338-47.
2. Ahmed HS, Shivananda DB, Pulkurthi SR, Dias AF, Sahoo PP. Clinical profile and outcomes in Tolosa-Hunt syndrome; a systematic review. J Clin Neurosci. 2024;129:110858.
3. İlgen Uslu F, Özkan M. Painful ophthalmoplegia: a case report and literature review. Agri. 2015;27:219-23.
4. La Mantia L, Erbetta A, Bussone G. Painful ophthalmoplegia: an unresolved clinical problem. Neurol Sci. 2005;26 Suppl 2:s79-82.
5. Gladstone JP, Dodick DW. Painful ophthalmoplegia: overview with a focus on Tolosa-Hunt syndrome. Curr Pain Headache Rep. 2004;8:321-9.
6. Moussa G, Bassilious K, Mathews N. A novel excel sheet conversion tool from Snellen fraction to LogMAR including 'counting fingers', 'hand movement', 'light perception' and 'no light perception' and focused review of literature of low visual acuity reference values. Acta Ophthalmol. 2021;99:e963-5.
7. Kline LB, Hoyt WF. The Tolosa-Hunt syndrome. J Neurol Neurosurg Psychiatry. 2001;71:577-82.
8. Park HS, Chung MS, Shin DS, Jung YW, Park JS. Whole courses of the oculomotor, trochlear, and abducens nerves, identified in sectioned images and surface models. Anat Rec (Hoboken). 2015;298:436-43.
9. Cleary JO, Sivarasan N, Burd C, Connor SEJ. Head and neck manifestations of granulomatosis with polyangiitis. Br J Radiol. 2021;94:20200914.
10. Housley SB, McPheeters MJ, Raygor KP, Bouslama M, Scullen T, Davies JM. Cavernous sinus thrombosis. Neurosurg Clin N Am. 2024;35:305-10.
11. Badakere A, Patil-Chhablani P. Orbital apex syndrome: a review. Eye Brain. 2019;11:63-72.
12. Williams ZR. Carotid-cavernous fistulae: a review of clinical presentation, therapeutic options, and visual prognosis. Int Ophthalmol Clin. 2018;58:271-94.
13. Ide S, Kiyosue H, Tokuyama K, Hori Y, Sagara Y, Kubo T. Direct carotid cavernous fistulas. J Neuroendovasc Ther. 2020;14:583-592.
14. Gemmete JJ, Chaudhary N, Pandey A, Ansari S. Treatment of carotid cavernous fistulas. Curr Treat Options Neurol. 2010;12:43-53.
15. Barrow DL, Spector RH, Braun IF, Landman JA, Tindall SC, Tindall GT. Classification and treatment of spontaneous carotid-cavernous sinus fistulas. J Neurosurg. 1985;62:248-56.
16. Fadda GL, Martino F, Andreani G, et al. Definition and management of invasive fungal rhinosinusitis: a single-centre retrospective study. Acta Otorhinolaryngol Ital. 2021;41:43-50.
17. Coutel M, Duprez T, Huart C, Wacheul E, Boschi A. Invasive fungal sinusitis with ophthalmological complications: case series and review of the literature. Neuroophthalmology. 2020;45:193-204.
18. Freddi TAL, Ottaiano AC, Lucio LL, Corrêa DG, Hygino da Cruz LC Jr. The trigeminal nerve: anatomy and pathology. Semin Ultrasound CT MR. 2022;43:403-13.

19. Alexander M, Cho E, Gliozheni E, Salem Y, Cheung J, Ichii H. Pathology of diabetes-induced immune dysfunction. Int J Mol Sci. 2024;25:7105.
20. Nyunt TPK, Mullol J, Snidvongs K. Predictive factors for invasive fungal rhinosinusitis in diabetic patients: Systematic review and data re-analysis. Asian Pac J Allergy Immunol. 2021;39:1-8.
21. Kulbay M, Tanya SM, Tuli N, et al. A Comprehensive review of thyroid eye disease pathogenesis: from immune dysregulations to novel diagnostic and therapeutic approaches. Int J Mol Sci. 2024;25:11628.
22. Muñoz-Ortiz J, Sierra-Cote MC, Zapata-Bravo E, et al. Prevalence of hyperthyroidism, hypothyroidism, and euthyroidism in thyroid eye disease: a systematic review of the literature. Syst Rev. 2020;9:201.

DOI: 10.4274/forbes.galenos.2026.43815
Forbes J Med 2026;7:53-59

Original Article

Isolated vs. Concomitant ACL Reconstruction in Meniscus Repair: Failure and Clinical Outcomes

Uğur ÖZDEMİR[1], Abdulhalim AKAR[2]

[1]Bandırma Onyedi Eylül University Faculty of Medicine, Department of Orthopaedics and Traumatology, Balıkesir, Türkiye
[2]Şişli Memorial Hospital, Clinic of Orthopaedics and Traumatology, İstanbul, Türkiye

Cite as: Özdemir U, Akar A. Isolated vs. concomitant ACL reconstruction in meniscus repair: failure and clinical outcomes. Forbes J Med. 2026;7:53-59

ABSTRACT

Objective: The healing rates between isolated meniscus repair and meniscus repair performed concomitantly with anterior cruciate ligament reconstruction (ACLR) are still a matter of debate. This study aimed to compare the failure rates and clinical outcomes between isolated meniscus repair and simultaneous ACLR.

Methods: This retrospective analysis included 59 patients aged <40 years who underwent arthroscopic meniscus repair. Of these, 32 patients underwent isolated meniscal repair, whereas 27 underwent simultaneous ACLR. Patients were followed at 1, 3, and 6 months postoperatively, and thereafter at 6-month intervals for a minimum of 2 years.

Results: No statistically significant differences were observed regarding tear type, tear length, or the affected meniscus. Although the distribution of tear zones was greater in the isolated group, favoring zone 2, this difference was not statistically significant (p=0.095). The preoperative Lachman test was significantly more frequently positive in patients who underwent simultaneous ACLR (p<0.001). When functional scores were evaluated, the preoperative (p=0.006) and postoperative (p=0.032) Tegner activity scores were significantly higher in the simultaneous ACLR group. The revision surgery rate, an indicator of clinical failure, was higher in isolated group; however, this difference did not reach statistical significance(p=0.060).

Conclusion: This study demonstrated that the clinical efficacy of meniscus repair is high regardless of whether ACLR is performed, but ACLR may show more favorable trends in some clinical parameters because of its biological and mechanical advantages.

Keywords: Anterior cruciate ligament (ACL), meniscus injury, repair technique

Received: 26.02.2026
Accepted: 22.05.2026
Epub: 02.06.2026
Publication Date: 11.06.2026

Corresponding Author:
Uğur ÖZDEMİR
Bandırma Onyedi Eylül University Faculty of Medicine, Department of Orthopaedics and Traumatology, Balıkesir, Türkiye
ugurozdemir.md@gmail.com
ORCID ID: 0000-0003-2889-8118

INTRODUCTION

The menisci are fibrocartilaginous structures that enhance joint congruence and play a key role in load transmission and shock absorption in the knee joint. With increasing participation in sports, the incidence of meniscal injury has increased. Loss of meniscal integrity is associated with knee instability[1] and early degenerative changes.[2-4]

Especially with advancements in arthroscopic techniques and implant technology, meniscus repair is increasingly used in clinical practice to restore the meniscus's joint-protective function and prevent the development of secondary lesions and joint degeneration.[3] Among the various repair methods, the inside-out and all-inside techniques are most commonly used and have demonstrated comparable healing outcomes and long-term failure rates.[5] Repair failure, often assessed by reoperation rates, occurs in approximately 23% of cases and is influenced by factors such as patient age, tear characteristics, vascularity, and chronicity.[6]

Meniscus tears resulting from sports injuries are frequently associated with anterior cruciate ligament (ACL) injuries and significantly impact patient outcomes. Biomechanical data and early clinical reports suggest that isolated meniscus repair is unsuitable for knees with ACL insufficiency. Whether healing rates differ between isolated and simultaneous anterior cruciate ligament reconstruction (ACLR) of the

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

meniscus remains a subject of debate. The primary aim of this study was to compare the clinical failure rates of meniscus repairs performed as isolated procedures with those performed simultaneously with ACLR. The secondary aim was to evaluate whether there are differences in clinical failure rates between meniscus repair techniques (inside-out and all-inside) and to examine the effect of ACLR on the relationship between the repair technique and clinical failure.

METHODS

Patient Selection

All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation, both institutional and national, and with the Declaration of Helsinki of 1975, as revised in 2008. Ethics Committee Approval was granted by the Bandırma Onyedi Eylül University Ethics Committee on 24.11.2025, with protocol number E-67961857-050.04-2500075936, a retrospective study was conducted between June 2022 and January 2025. Informed consent was waived because the study was retrospective. The study included 59 patients aged <40 years who were diagnosed with meniscal tears and had a normal mechanical axis. Horizontal and longitudinal tears greater than 1 cm that were observed in zones 1 and 2 and were not chronic degenerative were defined as criteria for repair. A total of 59 patients (30 women and 29 men) underwent arthroscopic meniscus repair. Thirty patients underwent the inside-out technique, while 29 underwent the all-inside repair technique. Thirty-two patients underwent isolated meniscus repair, and 27 underwent simultaneous ACLR. Patients were followed up at 1, 3, and 6 months postoperatively and then at 6-month intervals for a minimum of 2 years.

Patients with meniscal root tears, ramp lesions, multiple ligament injuries, concomitant microfractures or mosaicplasty due to cartilage damage, knee pathologies such as osteoarthritis, avascular necrosis, osteochondritis dissecans, or rheumatoid arthritis, a history of smoking, skeletally immature patients, previous ipsilateral knee surgery, and long bone alignment deformity were excluded.

Clinical Evaluation

Each patient was clinically evaluated for the presence of knee effusion, meniscal injury, and ACL injury, and all findings were recorded preoperatively and postoperatively. The Lachman test was used to assess ACL function, and the McMurray test was used to assess meniscal pathologies. Visual analog score (VAS), Lysholm, and Tegner activity scores were used in the preoperative and postoperative clinical evaluations of the patients. In our study, clinical failure was defined as the requirement for revision surgery. Revision surgery criteria were defined as a postoperative VAS score >3, the presence of knee effusion, and positive McMurray test results. Meniscus repair success was compared between groups that underwent isolated meniscus repair and those who underwent simultaneous ACLR. In addition, variables such as age, side, time between symptom onset and surgery, follow-up period, and duration of revision surgery were recorded for all patients.

Radiological Evaluation

All patients underwent preoperative knee radiography (anteroposterior and lateral views in the standing position), lower extremity orthoradiography, and magnetic resonance imaging (MRI). Preoperative knee radiographs were used to assess additional pathologies, such as osteoarthritis and cartilage damage. In patients who underwent ACLR, tunnel placement and implant position were assessed using knee radiographs obtained at the final follow-up. The preoperative mechanical axis of each patient was evaluated using lower extremity radiography. Meniscal and ACL injuries were diagnosed based on the evaluation of physical examination findings and preoperative MRI.

Surgical Technique

All surgical procedures were performed under a tourniquet using a standard knee arthroscopy approach. Each patient was evaluated arthroscopically for additional joint pathologies. In meniscus tears that met the repair criteria, the tear ends were debrided with a shaver to increase the healing potential. In patients who underwent simultaneous ACLR, meniscal repair was performed after harvesting gracilis and semitendinosus tendon autografts required for ACLR and creating bone tunnels. Following meniscus repair, the tendon autograft loaded onto the Endobutton system was passed through the prepared tunnels and fixed in the tibial tunnel with a bioabsorbable screw to complete the ACLR (ACL-PCL set; Doratek Medical, Ostim/Ankara, Türkiye). In the all-inside group, meniscus repair was performed using the TrustKnot All-Inside Meniscal Repair System (Doratek Medical, Ostim/Ankara, Türkiye) according to the manufacturer's technical guide. During the knot placement procedure, a knot pusher was used to ensure sufficient compression of the tear area. In the inside-out group, meniscal needles carrying reinforced threads were passed through the meniscus using cannulas (Inside-Out Meniscal Repair System; Doratek Medical, Ostim/Ankara, Türkiye). An additional surgical incision was made at the needle exit site, and the threads were knotted to the capsule. Meniscus repair was performed without damaging the knee joint cartilage. Depending on the length and stability of the tear,

1-4 implants were used. For each patient, data such as the location of the tear in the meniscus, zone, length, type, and number of sutures used for repair were recorded. The length and zone of the tear were assessed using distance-scaled probes. The zone of the tear (distance from the periphery) was defined according to the Cooper zone classification system.

Rehabilitation Protocol

All patients were referred to physiotherapists. Isometric quadriceps-strengthening exercises, such as quadriceps contractions and straight-leg raises, were initiated on day 1 post-surgery. The patients wore an angle-adjustable knee brace for 4 weeks, allowing for full extension and 90° flexion. Patients were instructed on the use of crutches and allowed to walk without bearing weight on the operated side for 4 weeks postoperatively. At the end of the 4th week, the angle-adjustable knee brace was discontinued and range-of-motion exercises were initiated. Simultaneously, they were allowed to bear weight at a level they could tolerate, and weight bearing was gradually increased. Full weight-bearing was achieved at the end of the 6th week, and crutches were discontinued at the end of the 8th week. Squatting was not permitted for the first three months after surgery. At 4 months post-surgery, if patients had achieved a functional range of motion, muscle strength, and stability, light sports activities were initiated. Patients were not allowed to return to normal sports activities for 6 months. Simultaneous ACLR did not alter the rehabilitation protocols.

Statistical Analysis

Statistical analysis of the data obtained in the study was performed using IBM Statistical Package for the Social Sciences (SPSS) for macOS version 30.0 (IBM Corp., Armonk, NY, USA). Categorical variables are presented as numbers and percentages [n (%)], while continuous variables are presented as mean ± standard deviation. The distribution characteristics of continuous variables were evaluated using the Kolmogorov-Smirnov normality test. In addition, skewness and kurtosis values within the ±2 range were considered to support the normality assumption. For intergroup comparisons, the Student's t-test was used for continuous variables showing normal distribution, and Welch's t-test was used when homogeneity of variance was not achieved. Pearson's chi-square test was used to compare categorical variables; Fisher's exact test was preferred when the expected cell counts were insufficient. Odds ratios and 95% confidence intervals were calculated to estimate the effect sizes associated with clinical failure (revision surgery). Results were considered statistically significant when p was less than 0.05. A post-hoc power analysis, based on the primary outcome (clinical failure or revision surgery), was performed to assess the adequacy of the sample size.

RESULTS

Table 1 summarizes the demographic and surgical characteristics according to the surgical methods. Of the 59 patients included, 32 (54.2%) underwent isolated meniscal repair, and 27 (45.8%) underwent simultaneous ACL reconstruction. The mean postoperative follow-up duration was 29.4 months (range, 24-31 months), and the mean time from symptom onset to surgery was 7.3 weeks (range, 1-12 weeks). No significant differences were observed between the groups regarding age, sex, operated side, follow-up duration, or time from injury to surgery ($p>0.05$). The distribution of meniscal repair techniques (all-inside vs. inside-out) was similar between the groups ($p=0.887$). Likewise, tear type, tear length, affected meniscus, tear zone distribution, and the number of sutures used did not differ significantly between the groups ($p>0.05$ for all).

Table 2 presents the clinical evaluations and functional outcomes. Postoperative effusion and positivity of the McMurray test were more frequent in the isolated group; however, the differences were not statistically significant ($p>0.05$). In patients with minimal effusion, the postoperative McMurray test became negative and VAS scores decreased to <3; therefore, no revision surgery was required. In contrast, 8 patients (13.6%) with severe effusion and a positive postoperative McMurray test were considered to have experienced clinical failure. Of these, 7 patients (21.9%) were in the isolated repair group, whereas 1 patient (3.7%) was in the concomitant ACLR group. Although revision surgery rates were higher in the isolated group, this difference was not statistically significant ($p=0.060$). These patients underwent arthroscopic partial meniscectomy after a mean of 8 months, and complete resolution of symptoms was observed at the final follow-up.

Preoperative Lachman test positivity was observed in 23 patients and was significantly higher among those who underwent simultaneous ACLR ($p<0.001$). Postoperatively, no residual anterior instability was identified on Lachman testing. The simultaneous ACLR group also demonstrated significantly higher preoperative and postoperative Tegner activity scores ($p=0.006$ and $p=0.032$, respectively); however, the amount of improvement (ΔTegner) did not differ between the groups ($p=0.317$). Both groups showed significant postoperative improvements in Lysholm and VAS scores, with no significant differences in postoperative values or in change scores (ΔLysholm and ΔVAS) ($p>0.05$).

Table 1. Demographic and surgical characteristics				
Variables (n=59)	**Total**	**Isolated (n=32)**	**Concomitant ACLR (n=27)**	**p value**
	n (%) or mean ± SD	**n (%) or mean ± SD**	**n (%) or mean ± SD**	
Gender				0.703
Male	29 (49.2)	15 (46.9)	14 (51.9)	
Female	30 (50.8)	17 (53.1)	13 (48.1)	
Age (years)	26±6	26±6	25±5	0.879
Side				0.957
Right	26 (44.1)	14 (43.8)	12 (44.4)	
Left	33 (55.9)	18 (56.3)	15 (55.6)	
Follow-up period	29.4±4.1	29.1±4	29.7±4.4	0.573
Time between tear and surgery (weeks)	7.3±3.3	7.8±3.5	6.6±3	0.135
Technique used				0.887
All-inside	29 (49.2)	16 (50)	13 (48.1)	
Inside-out	30 (50.8)	16 (50)	14 (51.9)	
Tear type				0.816
Horizontal	25 (42.4)	14 (43.8)	11 (40.7)	
Longitudinal	34 (57.6)	18 (56.3)	16 (59.3)	
Tear length (mm)	18.7±5.3	19.1±5.4	18.2±5.4	0.552
Which meniscus				0.739
Lateral	21 (35.6)	12 (37.5)	9 (33.3)	
Medial	38 (64.4)	20 (62.5)	18 (66.7)	
Tear zone				0.095
1	28 (47.5)	12 (37.5)	16 (59.3)	
2	31 (52.5)	20 (62.5)	11 (40.7)	
Number of stitches	2±0.7	2±0.8	1.9±0.7	0.468
SD: Standard deviation, ACLR: Anterior cruciate ligament reconstruction				

Table 3 presents the effect size analysis for clinical failure. Clinical failure rates were higher in the isolated group; however, this difference was not statistically significant. Although the effect size analysis suggested a trend toward a higher probability of clinical failure in the isolated group, this increase did not reach statistical significance or effect on clinical failure rates.

DISCUSSION

When evaluating the baseline characteristics and surgical details of the patient groups in our study, it can be stated that the two groups were well-balanced for clinical and surgical comparison. Most studies in which meniscus repair with simultaneous ACLR is performed report no significant differences between groups in terms of age and sex; however, such homogeneity is not always achieved for tear characteristics and zone distribution.[7] In the present study, no differences were found between groups in tear characteristics or zone distribution, allowing a more accurate and reliable assessment of repair success.

No statistically significant difference was observed between the groups regarding tear zone distribution (p=0.095). However, the numerically higher frequency of zone 2 tears in the isolated repair group and the p value's proximity to the significance threshold were considered a borderline trend. No significant difference in the number of sutures indicates that tears of similar length and complexity were repaired in both groups, thereby increasing the comparability of the results.

In this study, the all-inside and inside-out techniques were used in similar proportions in both groups, with no significant difference. In the literature, the choice of repair technique is mainly influenced by tear localization, surgeon preference, and concomitant ACLR.[8] The similarity in the distribution of repair techniques across both groups suggests that surgeons' preferences are independent of ACLR. This is important for interpreting the results because, in some studies, it is still debated whether all-inside techniques are used more frequently with ACLR

Table 2. Clinical evaluations and functional outcomes				
Variables (n=59)	**Total**	**Isolated (n=32)**	**Concomitant ACLR (n=27)**	**p value**
	n (%) or mean ± SD	**n (%) or mean ± SD**	**n (%) or mean ± SD**	
Pre-op effusion	59 (100)	32 (100)	27 (100)	NA
Post-op effusion				0.127
None	47 (79.7)	23 (71.9)	24 (88.9)	
Mild	4 (6.8)	2 (6.3)	2 (7.4)	
Severe	8 (13.6)	7 (21.9)	1 (3.7)	
Pre-op McMurray (+)	59 (100)	32 (100)	27 (100)	NA
Post-op McMurray (+)	8 (13.6)	7 (21.9)	1 (3.7)	0.060
Pre-op Lachman	23 (39)	**0 (0)**	**23 (85.2)**	**<0.001**
Post-op Lachman	0 (0)	0 (0)	0 (0)	NA
Pre-op Tegner	4±1	**3.7±1**	**4.4±0.8**	**0.006**
Post-op Tegner	7±2	**6.5±2**	**7.6±1.8**	**0.032**
ΔTegner	3.0±1.5	2.8±1.6	3.2±1.5	0.317
Pre-op Lysholm	55.5±11.8	55.1±11.6	55.9±12.3	0.790
Post-op Lysholm	83.3±13.9	80.8±16.7	86.2±9.3	0.125
ΔLysholm	27.8±14.9	25.7±16.2	30.3±13	0.246
Pre-op VAS	5.9±1.6	5.8±1.7	5.9±1.5	0.785
Post-op VAS	1.7±2.1	2.2±2.5	1.2±1.4	0.063
ΔVAS	-4.1±2.7	-3.6±2.9	-4.7±2.3	0.126
Revision surgery	8 (13.6)	7 (21.9)	1 (3.7)	0.060
Time until revision (months)	7.9±5.5	8.6±5.6	3±0	0.387

NA: Not applicable; Δ values are calculated as preoperative–postoperative difference. SD: Standard deviation, ACLR: Anterior cruciate ligament reconstruction

Table 3. Effect size analyses for clinical failure			
	Revision surgery n (%)	**OR (95% CI)**	**p value**
Surgical method			0.060
Isolated	7 (21.9)	7.30 (0.83-62.50)	
Simultaneous ACLR	1 (3.7)	Reference	
Technique used			1.000
All-inside	4 (13.8)	0.96 (0.22-4.27)	
Inside-out	4 (13.3)	Reference	

OR: Odds ratio, CI: Confidence interval. Reference groups were determined using simultaneous ACLR and inside-out techniques. p values were calculated using Fisher's exact test. ACLR: Anterior cruciate ligament reconstruction

due to shorter operating time, less soft-tissue dissection, and potentially different complication profiles, which may affect the results of the study.[9]

In our study, the similar postoperative clinical findings and functional outcomes between the two groups indicated that both groups generally exhibited comparable healing profiles. Although effusion and McMurray test positivity were higher in the isolated group, these differences were not statistically significant, which is consistent with the literature. Many studies have reported that meniscus repair with simultaneous ACLR supports biological healing by increasing intra-articular blood flow; therefore, less tenderness and mechanical symptoms may be observed in the early period.[10] Shelbourne and Gray[11] reported lower early postoperative effusion rates with ACLR, whereas Paxton et al.[12] noted a higher frequency of early mechanical symptoms after isolated meniscus repair. In this context, the difference in our study, which was not statistically significant but showed a clinical trend, is consistent with the biological healing advantage theory.

The higher postoperative Tegner scores in the ACLR group were consistent with the higher activity level. However, the similarity in the ΔTegner values suggests that the two groups were equivalent in terms of recovery. Nepple et al.[13] reported that patients who underwent meniscus repair in conjunction with ACLR returned to higher activity levels, but their functional recovery rates were similar to those with isolated meniscus repair.

The significant postoperative improvement in Lysholm and VAS scores in both groups strongly supports the clinical effectiveness of meniscus repair. The lack of a significant difference between groups has been frequently reported in the literature. Noyes and Barber-Westin[14] emphasized that the functional outcomes of meniscus repair are good, regardless of ACLR. Chahla et al.[15] showed that pain and function scores were similar in the long-term follow-up of patients who underwent isolated meniscus repair and repair with ACLR. Choi et al.[16] reported that ACLR did not significantly alter pain scores after meniscal repair.

This study showed that patients who underwent isolated meniscus repair had a higher rate of clinical failure, but this difference did not reach statistical significance. This finding is consistent with trends reported in the literature. Many studies have shown lower revision rates in meniscus repairs concurrent with ACLR. This can be explained by an improved biological healing environment, more stable knee biomechanics, and a younger, less degenerative tear profile that responds better to repair. For example, Nepple et al.[13] reported failure rates of 20-25% for isolated meniscus repairs and 10-12% for repairs with concomitant ACLR. Similarly, Paxton et al.[12] showed that the failure rate of repair was significantly lower with ACLR. Although the long-term results of isolated meniscus repairs are well defined in the literature, better clinical outcomes have been reported in patients who undergo concurrent ACLR. Westermann et al.[17] reported a 14% failure rate for meniscus repairs performed with ACLR at the 6-year follow-up. The lack of significance in our study is likely due to the small sample size. The p=0.060 value suggests that the statistical power hampered the clinically significant trend.

The lack of a significant difference in clinical failure between the all-inside and inside-out techniques is consistent with the results of large-scale meta-analyses. Various studies have shown that technique choice depends on factors such as tear localization, surgeon's experience, and concomitant pathologies, but is not a determining factor in long-term clinical success. For example, Chahla et al.[15] reported that the success rates of the all-inside and inside-out techniques were very similar, but their complication profiles differed. A systematic review by Grant et al.[18] found no significant difference in success between the two techniques; however, the inside-out technique provided stronger fixation, whereas the all-inside technique required less soft tissue dissection. The effect size analysis in our study also supports this finding, confirming that the choice of technique is not a determining factor in clinical failure.

Study Limitations

One of the main limitations of this study is the relatively small sample size, which may have limited the ability to detect statistically significant differences between groups. Although the revision rates were numerically higher in the isolated repair group and consistent with previous literature, this difference did not reach statistical significance. A post-hoc power analysis based on the primary outcome (clinical failure/revision surgery) demonstrated that the study was underpowered (61.1% power at α=0.05) that a larger sample size (approximately 92-93 patients) would be required to achieve adequate statistical power (80%). Therefore, the lack of statistically significant differences between groups should be interpreted with caution.

Nevertheless, strict inclusion and exclusion criteria were applied to establish a homogeneous patient population, thereby minimizing the influence of potential confounding variables. This methodological approach strengthens the internal validity of the study and supports a more reliable comparison between the groups despite the limited sample size.

CONCLUSION

This study demonstrated that the clinical efficacy of meniscus repair is high, regardless of the presence of ACLR; however, ACLR may show more favorable trends in some clinical parameters owing to its biological and mechanical advantages. Our findings are largely consistent with the existing literature and are supported by studies with larger sample sizes, particularly regarding revision rates and postoperative functional scores. Therefore, our results provide a framework consistent with the literature on biomechanics, biology, and clinical aspects. Further studies with larger sample sizes will help confirm the failure trend, especially in the isolated repair group.

Ethics

Ethics Committee Approval: All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation, both institutional and national, and with the Declaration of Helsinki of 1975, as revised in 2008. Ethics Committee Approval was granted by the Bandırma Onyedi Eylül University Ethics Committee on 24.11.2025, with protocol number E-67961857-050.04-2500075936.

Informed Consent: Informed consent was waived because the study was retrospective.

Footnotes

Authorship Contributions: Surgical and Medical Practices: U.Ö., A.A., Concept: U.Ö., A.A., Design: U.Ö., A.A.,

Data Collection or Processing: U.Ö., A.A., Analysis or Interpretation: U.Ö., A.A., Literature Search: U.Ö., A.A., Writing: U.Ö., A.A.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Gerhold C, Dave U, Bi AS, Chahla J. Medial meniscus root tears: anatomy, repair options, and outcomes. Arthroscopy. 2025;41:871-3.
2. Mihoubi F, Tordjman M, Khaled W, et al. Overview of meniscal injuries. Semin Musculoskelet Radiol. 2025;29:403-16.
3. Kuczyński N, Boś J, Białoskórska K, et al. The meniscus: basic science and therapeutic approaches. J Clin Med. 2025;14:2020.
4. Vaajala M, Teuho A, Liukkonen R, et al. Adolescence risk factors for meniscus and ligamentous knee injuries in adulthood: a longitudinal study. Knee Surg Sports Traumatol Arthrosc. 2026;34:1245-56.
5. Super JT, LaPrade RF, Robinson J, et al. Biologic augmentation of meniscus repair: a scoping review. Knee Surg Sports Traumatol Arthrosc. 2025;33:4176-88.
6. Farinelli L, Meena A, Montini D, et al. Failure rate of isolated medial meniscus repair in the stable knee: Systematic review and meta-analysis. Knee Surg Sports Traumatol Arthrosc. 2025;33:1333-44.
7. Sarraj M, Coughlin RP, Solow M, et al. Anterior cruciate ligament reconstruction with concomitant meniscal surgery: a systematic review and meta-analysis of outcomes. Knee Surg Sports Traumatol Arthrosc. 2019;27:3441-52.
8. Miller AC, Batiste AJ, McQuivey KS, McCarty EC. All-inside versus inside-out suture techniques in athletes undergoing arthroscopic meniscal repair: a systematic review and meta-analysis. Orthop J Sports Med. 2025;13:23259671251361488.
9. Robinson JD Jr, Williamson T, Carson T, Whelan RJ, Abelow SP, Gilmer BB. Primary anterior cruciate ligament repair: Current concepts. J ISAKOS. 2023;8:456-66.
10. Cho Y, Ahn JH. Comparison of the healing rate with meniscal repair concomitant with anterior cruciate ligament reconstruction and isolated meniscal repair based on magnetic resonance imaging signal intensity. Asia Pac J Sports Med Arthrosc Rehabil Technol. 2025;42:74-80.
11. Shelbourne KD, Gray T. Results of anterior cruciate ligament reconstruction based on meniscus and articular cartilage status at the time of surgery. Five- to fifteen-year evaluations. Am J Sports Med. 2000;28:446-52.
12. Paxton ES, Stock MV, Brophy RH. Meniscal repair versus partial meniscectomy: a systematic review comparing reoperation rates and clinical outcomes. Arthroscopy. 2011;27:1275-88.
13. Nepple JJ, Dunn WR, Wright RW. Meniscal repair outcomes at greater than five years: a systematic literature review and meta-analysis. J Bone Joint Surg Am. 2012;94:2222-7.
14. Noyes FR, Barber-Westin SD. Treatment of meniscus tears during anterior cruciate ligament reconstruction. Arthroscopy. 2012;28:123-30.
15. Chahla J, Serra Cruz R, Cram TR, Dean CS, LaPrade RF. Inside-out meniscal repair: medial and lateral approach. Arthrosc Tech. 2016;5:163-8.
16. Choi NH, Yang BS, Shin YC, Victoroff BN. Lateral meniscus posterior root tear repair with anterior cruciate ligament reconstruction: high healing rate and significant clinical improvement at a mean 8-year follow-up. Orthop J Sports Med. 2026;14:23259671251400773.
17. Westermann RW, Wright RW, Spindler KP, Huston LJ; MOON Knee Group; Wolf BR. Meniscal repair with concurrent anterior cruciate ligament reconstruction: operative success and patient outcomes at 6-year follow-up. Am J Sports Med. 2014;42:2184-92.
18. Grant JA, Wilde J, Miller BS, Bedi A. Comparison of inside-out and all-inside techniques for the repair of isolated meniscal tears: a systematic review. Am J Sports Med. 2012;40:459-68.

DOI: 10.4274/forbes.galenos.2026.47450
Forbes J Med 2026;7:60-68

Original Article

Relationship Between Early Neurological Deterioration and Transcranial Doppler Parameters in Mechanical Thrombectomy

Tuğba GÜLER GUNT[1], Demet Funda BAŞ[2], Ufuk ŞENER[3], Ümit BELET[4]

[1]İzmir Democracy University, Buca Seyfi Demirsoy Training and Research Hospital, Department of Neurology, İzmir, Türkiye
[2]Biruni University Faculty of Medicine, Department of Neurology, İstanbul, Türkiye
[3]University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Clinic of Neurology, İzmir, Türkiye
[4]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Radiology, İzmir, Türkiye

Cite as: Güler Gunt T, Baş DF, Şener U, Belet Ü. Relationship between early neurological deterioration and transcranial doppler parameters in mechanical thrombectomy. Forbes J Med. 2026;7:60-68

ABSTRACT

Objective: Mechanical thrombectomy is an effective treatment method for stroke patients with large vessel occlusion. Hemorrhagic transformation and edema can be observed during patient follow-up after mechanical thrombectomy. Transcranial Doppler (TCD) is a method used to evaluate hemodynamic changes following endovascular treatment (EVT), and the examination of patients with TCD contributes to the assessment of early neurological deterioration.

Methods: Patients who presented to University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital with acute ischemic stroke and underwent mechanical thrombectomy for anterior circulation occlusion were included in the study. Demographic data were recorded, along with initial and 24th-hour National Institutes of Health Stroke Scale (NIHSS) scores, post-procedural revascularization scores, and whether intravenous thrombolysis was performed. Patients were evaluated using TCD within the first 12 hours after EVT, and parameters in the ipsilateral and contralateral vessels were compared to examine their impact on hemorrhagic transformation and edema. Clinical outcomes were assessed using the 3-month modified Rankin scale (mRS).

Results: The pulsatility index, peak systolic velocity, and end-diastolic velocity on TCD examination did not significantly affect hemorrhagic transformation and edema when comparing ipsilateral and contralateral vessels. A higher 24-hour NIHSS score was significantly associated with hemorrhagic transformation. Patients with edema were found to have higher mRS scores and an association with clinical outcomes.

Conclusion: In our study, the evaluation of early clinical response by TCD examination after mechanical thrombectomy did not demonstrate a significant contribution.

Keywords: Acute ischemic stroke, thrombectomy, transcranial doppler, pulsatility index, hemorrhagic transformation

Received: 13.01.2026
Accepted: 02.06.2026
Epub: 03.06.2026
Publication Date: 17.06.2026

Corresponding Author:
Tuğba GÜLER GUNT
İzmir Democracy University, Buca Seyfi Demirsoy Training and Research Hospital, Department of Neurology, İzmir, Türkiye
tubguler@gmail.com
ORCID ID: 0009-0001-1010-9275

This study was previously presented as a poster presentation at the European Academy of Neurology meeting in Finland in June 2025.

INTRODUCTION

Stroke is defined as a neurological deficit that occurs following an acute focal injury to the central nervous system due to a vascular event, and is one of the leading causes of morbidity and mortality worldwide.[1] Significant advances have been made in recent years in the diagnosis and treatment of acute stroke.

Mechanical thrombectomy is the most effective treatment for acute ischemic stroke and is recommended within the first 6 hours for patients with large vessel occlusion (LVO).[2,3] Recent studies have indicated that under suitable criteria, this procedure can be performed within 24 hours.[2,4,5] Although mechanical thrombectomy yields good functional outcomes, some patients may experience early neurological deterioration, particularly within the first 72 hours.[6] Hemorrhagic transformation and edema are complications associated with poor functional outcomes in acute stroke patients and managing cerebral

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

autoregulation in the early phase has become a target to prevent these complications.[7]

Transcranial Doppler (TCD) is an easily accessible, cost-effective and non-invasive diagnostic method used to evaluate hemodynamic changes following reperfusion therapies.[8,9] In the acute phase of stroke, TCD is used to assess vascular occlusion and cerebral perfusion, to detect microembolic signals, and to evaluate collateral circulation.[9] Parameters such as pulsatility index (PI), peak systolic flow, and end-diastolic flow in the vessel are examined using TCD; the PI, in particular, is associated with increased intracranial pressure.[6-10] In our study, we aimed to investigate the relationship between early neurological deterioration and TCD findings in patients undergoing mechanical thrombectomy.

METHODS

Scope of the Study

Prospective

The study included 20 patients who were admitted to University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospita with a diagnosis of acute ischemic stroke between July 2022 and July 2023 and who met the inclusion criteria. Patients over 18 years of age who underwent mechanical thrombectomy for anterior circulation strokes were included (regardless of whether they received intravenous thrombolytic therapy). Patients were excluded from the study if they had no TCD window, had a final Thrombolysis in Cerebral Infarction (TICI) score of 0-1, or were referred to another center post-procedure.

Clinical Evaluation

This prospective study was conducted to evaluate the relationship between TCD data and early neurological response in patients admitted to University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospita with acute ischemic stroke and undergoing mechanical thrombectomy between July 2022 and July 2023.

During this period, patients with measurable neurological deficits who presented to University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospita with acute ischemic stroke were assessed, with National Institutes of Health Stroke Scale (NIHSS) scores calculated based on initial examination findings. In the acute phase, brain computed tomography (CT), with or without computer tomography angiography (CTA), was performed, and endovascular treatment (EVT) was applied according to guidelines for patients with anterior circulation LVO. The grade of revascularization was evaluated using the TICI score, with TICI 0, 1, 2a, and 2b defined as unsuccessful revascularization, and TICI 2c and 3 defined as successful revascularization. The Tan collateral grading system was used to evaluate the collateral system in CTA. According to this system, 0 indicates no collateral support, 1 corresponds to filling 50% or less of the occluded middle cerebral artery (MCA) region, 2 indicates filling between 50% and 100% of the occluded MCA region and 3 represents complete filling of the occluded region. For patients with successful revascularization, TCD was performed within the first 12 hours post-procedure, and the time of Doppler assessment was recorded. To ensure consistency of the results, TCD was performed by a single operator. All flow parameters in the affected vessel (peak systolic flow, end-diastolic flow, PI, resistance index) were calculated and compared with the contralateral vessel.

Data on patients' age, gender, smoking and alcohol use; chronic conditions such as hypertension, diabetes, atrial fibrillation, hyperlipidemia, coronary artery disease; and history of stroke or transient ischemic attack were recorded.

Post-procedure, patients' vital signs were measured in the intensive care unit and NIHSS scores were recorded at the 24 hours. For patients with clinical deterioration (a decrease of 4 or more points in the NIHSS score), non-contrast brain CT scans were obtained either immediately post-procedure or at 24 hours post-procedure to assess hemorrhagic transformation, ischemic progression, and edema.

During follow-up in the stroke outpatient clinic, patients' 3-month modified Rankin Scale (mRS) scores were recorded. An mRS score of 0-2 at 3 months was considered a good clinical outcome (functional independence); an mRS score of 0-3 was considered to represent ambulatory independence; and an mRS score of 3-6 was considered a poor clinical outcome.

Ethics

This study was conducted in accordance with the ethical guidelines approved by the Non-Interventional Research Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, with decision number 2022/06-31 dated June 15, 2022. Informed consent was obtained from all participating patients.

Statistical Analysis

Data were evaluated using IBM SPSS Statistics Standard Concurrent User V 26 (IBM Corp., Armonk, New York, USA). Descriptive statistics were presented as number of units (n), percentage (%), mean ± standard deviation, median, minimum, maximum and interquartile range values. For comparisons between groups regarding numerical variables, the independent samples t-test was used when

the data showed normal distribution, while the Mann-Whitney U test, employing the exact method for small sample comparisons, was used when the data did not show normal distribution. Chi-square analyses (Pearson chi-square, Yates chi-square, Fisher exact test) were used to compare categorical variables between groups. Multiple binary logistic regression analyses were used to identify factors affecting hemorrhagic transformation and edema. Variables with p<0.25 in univariate analyses were included in the logistic regression models. The Wald backward-elimination method was applied to select the final model. The goodness of fit for regression analyses was assessed using the Hosmer-Lemeshow test, while the percentage of explanation was evaluated with the Nagelkerke R^2 coefficient. A p value of <0.05 was considered statistically significant.

Table 1. Demographic and clinical variables of participants

Variables	n
Gender	
Female	6 (30%)
Male	14 (70%)
Age	62.8±13.5 (34-89)*
Smoking	13 (65%)
Alcohol	4 (20%)
Comorbidities	
Hypertension	10 (50%)
Diabetes mellitus	5 (25%)
Atrial fibrillation	7 (35%)
Cardiac diseases	11 (55%)
Hyperlipidemia	1 (5%)
Stroke	3 (15%)
Other	8 (40%)
Admission NIHSS	14.5 (8-21)
Admission time (min)	120 (6-360)
Iv-tPA	
Yes	14 (70%)
No	6 (30%)
CTA collateral score	1 (0-3)
TICI score	
2C	4 (20%)
3	16 (80%)
TCD evaluation time	6 (1-12)
24-hour NIHSS	7 (0-24)
3-months mRS	1 (0-6)
Progression of ischemia	3 (15%)
Hemorrhagic transformation	10 (50%)
Edema	6 (30%)

NIHSS: National Institutes of Health Stroke Scale, Iv-tPA: Intravenosus thrombolitic therapy, TICI: Thrombolysis in cerebral infarction, TCD: Transcranial Doppler, mRS: Modified Rankin scale, min: Minutes, CTA: Computer tomography angiography

n: Number of patients, %: Percentage, Numeric variables are summarized as median (minimum-maximum) values,

*mean ± standard deviation (minimum-maximum)

RESULTS

A total of 20 patients who underwent EVT for acute ischemic stroke and met the inclusion criteria were included in the study between July 2022 and July 2023 (Table 1).

After EVT, the vital signs in the intensive care unit showed a systolic blood pressure (SBP) of 131.5±19.1, a diastolic blood pressure (DBP) of 78.1±14.8, and a pulse of 80 (56.0-132.0). No statistically significant relationship was found between SBP, DBP, and mean arterial pressure (MAP) values and ipsilateral and contralateral flow parameters. The SBP, DBP, and MAP values in patients with hemorrhagic transformation and edema did not differ significantly.

Among those with progression of ischemia, 1 (33%) was female and 2 (66.7%) were male; their ages were significantly lower than those of individuals without progression. There was no statistically significant difference in the values of ipsilateral and contralateral flow parameters concerning progression of ischemia.

According to Table 2, The smoking rate in patients without hemorrhagic transformation was statistically higher than in those with it. The 24-hour NIHSS values in patients with hemorrhagic transformation were significantly higher than in patients without hemorrhagic transformation.

In the comparisons in Table 2, variables such as gender, smoking, alcohol, ipsilateral time-averaged maximum velocity (TAMAX) pulse, 24-hour NIHSS, 3-month mRS, which had p<0.25 were included in the multiple binary logistic regression analysis, as they were considered to potentially affect hemorrhagic transformation. The 24-hour NIHSS was found to be effective in predicting hemorrhagic transformation. For each unit increase in the 24-hour NIHSS, the risk of hemorrhagic transformation increased by 1.274 times [odds ratio (OR) confidence interval (CI) 1.002-1.619]. The relationship between hemorrhagic transformation and the elevation of ipsilateral flow parameters is shown in Table 3.

The 24-hour NIHSS values of patients with edema are significantly higher than those of patients without edema. In the comparisons presented in Table 4, variables such as smoking, cardiac diseases, stroke, CTA collateral score, TCD assessment time, ipsilateral PSV, TAMAX, PI, contralateral PSV, 24-hour NIHSS, and 3-month mRS, which have p<0.25 values were included in the multiple binary logistic regression analysis considering their potential effects on

Table 2. Comparisons regarding hemorrhagic transformation				
Variables	**Hemorrhagic transformation**		**Test statistics**	
	No n=10	**Yes n=10**	**Test value**	**p value**
Sociodemographic variables, n (%)				
Female	1 (10.0)	5 (50.0)	3.810	0.14‡
Male	9 (90.0)	5 (50.0)		
Age	60.5±13.9	65.1±13.4	0.750	0.463†
Smoking	10 (100.0)	3 (30.0)	**10.769**	**0.003‡**
Alcohol	4 (40.0)	0 (0.0)	5.000	0.087‡
Hypertension	5 (50.0)	5 (50.0)	-	-
Diabetes mellitus	3 (30.0)	2 (20.0)	0.267	>0.999‡
Atrial fibrillation	2 (20.0)	5 (50.0)	1.978	0.350‡
Cardiac diseases	5 (50.0)	6 (60.0)	0.202	>0.999‡
Hyperlipidemia	1 (10.0)	0 (0.0)	1.053	>0.999‡
Stroke	1 (10.0)	2 (20.0)	0.392	>0.999‡
Other	4 (40.0)	4 (40.0)	-	-
Clinical variables				
Admission NIHSS	14.0 (7.3)	15.0 (6.0)	0.988	0.353&
Admission time (min)	105.0 (165.0)	180.0 (225.0)	0.571	0.579&
IV-TPA, n (%)				
No	8 (80.0)	6 (60.0)	0.952	0.628‡
Yes	2 (20.0)	4 (40.0)		
CTA collateral score	1.0 (1.0)	1.5 (1.0)	0.124	0.912&
TICI score, n (%)				
2C	2 (20.0)	2 (20.0)	-	-
3	8 (80.0)	8 (80.0)		
TCD evaluation time	7.0 (9.3)	5.0 (8.8)	0.419	0.684&
Ipsilateral flow parameters				
PSV	94.8 (60.2)	120.4 (63.6)	0.680	0.529&
EDV	44.2 (33.1)	44.0 (27.8)	0.113	0.912&
TAMAX	60.4 (41.0)	89.4 (38.9)	1.285	0.218&
PI	0.98 (0.54)	0.92 (1.18)	0.151	0.912&
RI	0.60 (0.27)	0.59 (0.52)	0.076	0.971&
Contralateral flow parameters				
PSV	90.4 (47.8)	100.2 (31.6)	0.718	0.481&
EDV	38.4 (33.0)	44.9 (21.4)	0.151	0.912&
TAMAX	57.7 (39.1)	65.0 (22.8)	0.001	>0.999&
PI	0.97 (0.80)	0.96 (0.89)	0.151	0.912&
RI	0.60 (0.12)	0.59 (0.28)	0.001	>0.999&
Other variables				
24-hour NIHSS	1.5 (3.8)	9.0 (12.3)	**2.663**	**0.007&**
mRS 3-month	1.0 (3.0)	3.5 (5.3)	1.477	0.165&

NIHSS: National Institutes of Health Stroke Scale, Iv-tPA: Intravenosus thrombolitic therapy, TICI: Thrombolysis in cerebral ınfarction, TCD: Transcranial Doppler, mRS: Modified Rankin scale, PSV: Peak sistolic velocity, EDV: End-diastolic velocity, TAMAX: Time averaged mean velocity, PI: Pulsatility index, RI: Resistance index, min: Minutes, CTA: Computer tomography angiography

n: Number of patients, %: Percentage, numeric variables are summarized as median (minimum-maximum) values, †: Independent samples t-test, ‡: Chi-square test, &: Exact Mann-Whitney U test

Table 3. Relationship between hemorrhagic transformation and elevation of ipsilateral flow parameters				
Variables	**Hemorrhagic transformation**		**Test statistics**	
	No	Yes	Test value	p value
PSV, n (%)				
Low	4 (40.0)	5 (50.0)	0.202	>0.999
High	6 (60.0)	5 (50.0)		
EDV, n (%)				
Low	4 (40.0)	4 (40.0)	-	-
High	6 (60.0)	6 (60.0)		
TAMAX, n (%)				
Low	4 (40.0)	3 (30.0)	0.220	>0.999
High	6 (60.0)	7 (70.0)		
PI, n (%)				
Low	4 (40.0)	3 (30.0)	1.143	>0.999
Same	0 (0.0)	1 (10.0)		
High	6 (60.0)	6 (60.0)		
RI, n (%)				
Low	3 (30.0)	2 (20.0)	1.291	0.700
Same	1 (10.0)	3 (30.0)		
High	6 (60.0)	5 (50.0)		

PSV: Peak sistolic velocity, EDV: End-diastolic velocity, TAMAX: Time averaged mean velocity, PI: Pulsatility index, RI: Resistance index. Difference = Ipsilateral-contralateral, n: Number of patients, %: Column percentage

edema. Accordingly, the risk of edema increased 3.705-fold with each unit increase in the 3-month mRS (OR CI: 1.155-11.881). The relationship between edema and elevation of ipsilateral flow parameters was shown in Table 5.

DISCUSSION

Assessment of patients presenting with acute stroke who undergo mechanical thrombectomy, using TCD in the acute post-procedure period, can serve as a technique to predict early neurological deterioration.[11] Previous studies in this area have primarily focused on patients who received intravenous thrombolytic therapy and studies involving patients who underwent mechanical thrombectomy are limited in number.[12-14]

In our study of mechanical thrombectomy patients, individual evaluation using TCD showed no significant relationship between ipsilateral and contralateral parameters and early neurological response. These results may be attributable to the small number of patients included in the study. For example, in the study conducted by He et al.,[11] 112 patients underwent EVT. Successful EVT was achieved in 80 of these patients. Early neurological deterioration was observed in 17 of the 80 patients; in those patients, the ipsilateral mean flow velocity (MFV)/contralateral MFV ratio was higher.

The PI reflects cerebrovascular resistance and serves as an indirect indicator of intracranial pressure; increased PI indicates elevated intracranial pressure.[11,15,16] A high PI has been associated with higher mRS scores and is used to predict poor functional outcomes.[17] One study indicated that PI values reached their highest levels on day 3 and were associated with increased edema.[18] In our study, while the ipsilateral PI values of patients with edema were higher compared to contralateral PI values, this difference was not statistically significant. This may be attributable to the examination of patients with TCD within the first 12 hours, as this timing may be too early to assess edema and increased intracranial pressure. In the study by Pan et al.,[19] no significant difference in PI was found when comparing patients with or without edema, suggesting that ischemia and damage to the vascular endothelium could disrupt cerebral autoregulation, exacerbating the severity of ischemic injury and leading to edema. High PI values are associated with intracranial hemorrhage and serve as a determinant of poor clinical outcomes; however, as shown in another study and corroborated by our findings, no significant relationship was found between intracranial hemorrhage and PI.[8,15]

Peak systolic velocity is associated with clinical outcomes and has been shown to be elevated in patients with hemorrhagic transformation.[9,20] Previous studies have indicated that increases in blood pressure and PSV may lead to cerebral hyperperfusion.[11,20] Although the PSV on

Table 4. Comparisons regarding edema				
Variables	**Edema**		**Test statistics**	
	No n=14	Yes n=6	Test value	p value
Sociodemographics variables, n (%)				
Female	3 (21.4)	3 (50.0)	1.633	0.303‡
Male	11 (78.6)	3 (50.0)		
Age	61.2±13.5	66.3±14.1	0.754	0.460†
Smoking	11 (78.6)	2 (33.3)	3.778	0.122‡
Alcohol	4 (28.6)	0 (0.0)	2.143	0.267‡
Hypertension	7 (50.0)	3 (50.0)	-	-
Diabetes mellitus	4 (28.6)	1 (16.7)	0.317	>0.999‡
Atrial fibrillation	5 (35.7)	2 (33.3)	0.010	>0.999‡
Cardiac diseases	6 (42.9)	5 (83.3)	2.780	0.157‡
Hyperlipidemia	1 (7.1)	0 (0.0)	0.451	>0.999‡
Stroke	1 (7.1)	2 (33.3)	2.260	0.202‡
Clinical variables				
Admission NIHSS	14.5 (6.5)	14.5 (6.3)	0.290	0.779&
Admission time (min)	120.0 (180.0)	90.0 (253.5)	0.457	0.659&
IV-TPA, n (%)				
No	10 (71.4)	4 (66.7)	0.045	>0.999‡
Yes	4 (28.6)	2 (33.3)		
CTA collateral score	2.0 (1.0)	1.0 (0.5)	1.898	0.091&
TICI score, n (%)				
2C	2 (14.3)	2 (33.3)	0.952	0.549‡
3	12 (85.7)	4 (66.7)		
TCD evaluation time	10.0 (7.5)	1.5 (3.5)	**2.787**	**0.003&**
Ipsilateral flow parameters				
PSV	94.8 (60.6)	125.5 (63.5)	1.485	0.153&
EDV	45.2 (31.0)	42.7 (41.9)	0.701	0.494&
TAMAX	60.4 (40.8)	92.0 (52.9)	1.650	0.109&
PI	0.93 (0.67)	1.37 (1.30)	1.237	0.239&
RI	0.59 (0.33)	0.74 (0.50)	1.077	0.312&
Contralateral flow parameters				
PS	85.8 (39.5)	108.6 (19.8)	1.650	0.109&
ED	38.4 (28.6)	47.5 (26.0)	0.577	0.602&
TAMAX	60.0 (25.3)	72.4 (30.0)	0.495	0.659&
PI	0.98 (0.76)	0.89 (0.94)	0.248	0.841&
RI	0.60 (0.22)	0.57 (0.21)	0.580	0.602&
Other variables				
24-hour NIHSS	3.5 (3.3)	14.5 (10.5)	**3.237**	**<0.001&**
mRS 3-month	0.5 (1.5)	5.0 (3.0)	**3.139**	**0.001&**

NIHSS: National Institutes of Health Stroke Scale, Iv-tPA: Intravenosus thrombolitic therapy, TICI: Thrombolysis in cerebral infarction, TCD: Transcranial Doppler, mRS: Modified Rankin scale, PSV: Peak sistolic velocity, EDV: End-diastolic velocity, TAMAX: Time averaged mean velocity, PI: Pulsatility index, RI: Resistance index, min: Minutes, CTA: Computer tomography angiography

n: Number of patients, %: Percentage, Numeric variables are summarized as median (minimum-maximum) values, ‡: Chi-square test, &: Exact Mann-Whitney U test

Table 5. Relationship between edema and elevation of ipsilateral flow parameters

Variables	Edema		Test statistics	
	No	Yes	Test value	p value
PSV, n (%)				
Low	6 (42.9)	3 (50.0)	0.087	>0.999
High	8 (57.1)	3 (50.0)		
EDV, n (%)				
Low	6 (42.9)	2 (33.3)	0.159	>0.999
High	8 (57.1)	4 (66.7)		
TAMAX, n (%)				
Low	6 (42.9)	1 (16.7)	1.266	0.354
High	8 (57.1)	5 (83.3)		
PI, n (%)				
Low	6 (42.9)	1(16.7)		
Same	1 (7.1)	0 (0.0)	2.059	0.533
High	7 (50.0)	5 (83.3)		
RI, n (%)				
Low	4 (28.6)	1 (16.7)		
Same	4 (28.6)	0 (0.0)	3.203	0.257
High	6 (42.9)	5 (83.3)		

PSV: Peak sistolic velocity, EDV: End-diastolic velocity, TAMAX: Time averaged mean velocity, PI: Pulsatility index, RI: Resistance index. Difference = Ipsilateral-contralateral; n: Number of patients; %: Column percentage

the ipsilateral side was higher than on the contralateral side, no association with hemorrhagic transformation was demonstrated in our study. Additionally, the patients' blood pressure values were not associated with hemorrhagic transformation. Recording only a single measurement within the first 12 hours may have limited the ability to detect hyperperfusion. Furthermore, while our study found no difference between ipsilateral and contralateral sides in patients with edema, another study found that PSV on the affected side compared to the contralateral side was associated with an increased risk of edema.[19] In 185 cases with acute anterior circulation LVO, Baracchini et al.[21] assessed peak systolic velocity in the early phase at 48 h and 1 week following mechanical thrombectomy and found that a high ipsilateral/contralateral PSV ratio contributed to poor clinical outcomes but was not associated with edema. The lack of statistical significance in these results may be related to the small sample size, inappropriate timing of TCD assessments, and insufficient repetitions in our study.

Kneihsl et al.[22] reported in a cohort of 123 ischemic stroke patients that a ≥30% increase in MFV in the affected MCA within 72 hours after successful mechanical thrombectomy was associated with hemorrhagic transformation and poor 90-day clinical outcomes. In contrast, Perren et al.,[23] in a study of 31 patients evaluated within 7 days after recanalization, found that a 35-40% increase in MFV in the affected MCA compared with the contralateral side was not associated with clinical deterioration. This discrepancy may be attributed to differences in the timing of TCD assessment, with earlier measurements potentially providing more prognostic value. Therefore, in our study, performing a TCD evaluation within the first 12 hours after EVT is an appropriate time to predict early neurological deterioration.

Cerebral edema is a significant issue affecting mortality and morbidity in acute stroke patients. In our study, the average NIHSS score for patients with edema at 24 hours was found to be 14.5, consistent with the findings of Wu et al.,[24] who reported that high NIHSS values were associated with edema and that patients with an NIHSS score ≥15 were at risk for developing edema. Similarly, this study found that edema was associated with poor collateral circulation, and our findings support this: patients with edema had worse collateral circulation.[24] The HERMES meta-analysis indicated that a high NIHSS score creates a risk for edema, and that EVT and reperfusion are associated with better functional outcomes; however, reperfusion times were not related to edema in patients undergoing EVT.[25] The lack of recorded reperfusion times in our study may have limited the evaluation of the results. Consistent with the conclusions of the HERMES meta-analysis, our study found that patients with edema had poor clinical outcomes at three months.[25]

Hemorrhagic transformation is a complication that leads to neurological deterioration and poor functional outcomes, particularly after thrombolytic therapy in acute ischemic stroke.[26,27] Consistent with previous studies, our findings indicated that hemorrhagic transformation was associated with high NIHSS scores.[26,28] Hemorrhagic transformation is classified as symptomatic or asymptomatic intracranial hemorrhage. In the DEFUSE and DAWN studies, the rates of symptomatic ICH did not differ significantly between patients treated with or without EVT.[4,5] Tian et al.[26] found that a time of more than 280 minutes between stroke onset and revascularization was associated with an increased risk of hemorrhagic transformation. In our study, this time was recorded as 180 minutes, which is longer (by 105 minutes) compared to patients without hemorrhagic transformation, but not statistically significant. This may be related to the small sample size. Although previous studies have indicated that more than three passes with a stent retriever increase the risk of ICH, the absence of recorded data on the number of passes and on thrombectomy technique in our study limits our ability to comment further on this association.[26,29] According to our results, systolic

and mean blood pressure values were not associated with hemorrhagic transformation; previous studies have shown that high SBP and hemodynamic variability within the first 24 hours after EVT are associated with hemorrhagic transformation and poor clinical outcomes.[27,30,31] The failure to record serial blood pressure measurements in our study may have contributed to this issue.

The three-months mRS score is used to evaluate functional independence in acute ischemic stroke. The literature indicates that functional independence rates in patients undergoing mechanical thrombectomy range from 33% to 60%, and our study found a rate of 55%, which supports the existing literature.[32-34]

Study Limitations

The most significant limitation of our study is the small number of patients who underwent mechanical thrombectomy. Some patients could not be included in the assessment due to unsuitable TCD windows. Regarding the study design, evaluations within the first 12 hours after mechanical thrombectomy were not feasible in some patients. Mechanical thrombectomy was not assessed in our study; future studies could compare the effects of vascular endothelial damage on various indices and demonstrate vascular wall responses by examining patients who underwent stent-retriever and ADAPT techniques. Additionally, recording the time to revascularization and the number of passes during mechanical thrombectomy is necessary to provide insights into these aspects.

CONCLUSION

TCD is increasingly used in neurology practice and for the follow-up of patients who have undergone mechanical thrombectomy. In our study, early assessment with TCD of patients undergoing mechanical thrombectomy in the acute stroke setting did not contribute to an early clinical response. To clarify these findings, prospective studies with larger sample sizes are needed.

Ethics

Ethics Committee Approval: This study was conducted in accordance with the ethical guidelines approved by the Non-Interventional Research Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, with decision number 2022/06-31 dated June 15, 2022.

Informed Consent: Informed consent was obtained from all participating patients.

Footnotes

Authorship Contributions

Surgical and Medical Practices: U.Ş., Concept: T.G.G., Design: D.F.B., U.Ş., Data Collection or Processing: T.G.G., Ü.B., Analysis or Interpretation: D.F.B., Literature Search: T.G.G., Writing: T.G.G., D.F.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Campbell BCV, Khatri P. Stroke. Lancet. 2020;396:129-42.
2. Powers WJ, Rabinstein AA, Ackerson T, et al. 2018 Guidelines for the Early Management of Patients With Acute Ischemic Stroke: A Guideline for Healthcare Professionals From the American Heart Association/American Stroke Association. Stroke. 2018;49:e46-110.
3. Antipova D, Eadie L, Makin S, Shannon H, Wilson P, Macaden A. The use of transcranial ultrasound and clinical assessment to diagnose ischaemic stroke due to large vessel occlusion in remote and rural areas. PLoS One. 2020;15:e0239653.
4. Nogueira RG, Jadhav AP, Haussen DC, et al. Thrombectomy 6 to 24 Hours after stroke with a mismatch between deficit and infarct. N Engl J Med. 2018;378:11-21.
5. Albers GW, Marks MP, Kemp S, et al. Thrombectomy for stroke at 6 to 16 hours with selection by perfusion imaging. N Engl J Med. 2018;378:708-18.
6. Chen H, Su Y, He Y, et al. Controlling blood pressure under transcranial Doppler guidance after endovascular treatment in patients with acute ischemic stroke. Cerebrovasc Dis. 2020;49:160-9.
7. Castro P, Azevedo E, Serrador J, Rocha I, Sorond F. Hemorrhagic transformation and cerebral edema in acute ischemic stroke: Link to cerebral autoregulation. J Neurol Sci. 2017;372:256-61.
8. Park J, Hwang SK. Transcranial Doppler study in acute spontaneous intracerebral hemorrhage: the role of pulsatility index. J Cerebrovasc Endovasc Neurosurg. 2021;23:334-42.
9. Shahripour RB, Azarpazhooh MR, Akhuanzada H, et al. Transcranial Doppler to evaluate postreperfusion therapy following acute ischemic stroke: a literature review. J Neuroimaging. 2021;31:849-57.
10. Blanco P, Abdo-Cuza A. Transcranial Doppler ultrasound in neurocritical care. J Ultrasound. 2018;21:1-16.
11. He YB, Su YY, Rajah GB, et al. Trans-cranial Doppler predicts early neurologic deterioration in anterior circulation ischemic stroke after successful endovascular treatment. Chin Med J (Engl). 2020;133:1655-61.
12. Baizabal-Carvallo JF, Alonso-Juarez M, Samson Y. Clinical deterioration following middle cerebral artery hemodynamic changes after intravenous thrombolysis for acute ischemic stroke. J Stroke Cerebrovasc Dis. 2014;23:254-8.

13. Lam MY, Haunton VJ, Panerai RB, Robinson TG. Cerebral hemodynamics in stroke thrombolysis (CHiST) study. PLoS One. 2020;15:e0238620.
14. Nogueira RC, Lam MY, Llwyd O, et al. Cerebral autoregulation and response to intravenous thrombolysis for acute ischemic stroke. Sci Rep. 2020;10:10554.
15. Aoki J, Raber LN, Katzan IL, Hussain MS, Hui FK, Uchino K. Post-intervention TCD examination may be useful to predict outcome in acute ischemic stroke patients with successful intra-arterial intervention. J Neurol Sci. 2013;334:26-9.
16. Bill O, Lambrou D, Sotomayor GT, et al. Predictors of the pulsatility index in the middle cerebral artery of acute stroke patients. Sci Rep. 2020;10:17110.
17. Zhao W, Liu R, Yu W, et al. Elevated pulsatility index is associated with poor functional outcome in stroke patients treated with thrombectomy: a retrospective cohort study. CNS Neurosci Ther. 2022;28:1568-75.
18. Asil T, Uzunca I, Utku U, Berberoglu U. Monitoring of increased intracranial pressure resulting from cerebral edema with transcranial Doppler sonography in patients with middle cerebral artery infarction. J Ultrasound Med. 2003;22:1049-53.
19. Pan J, Wu H, Wu T, Geng Y, Yuan R. Association between post-procedure cerebral blood flow velocity and severity of brain edema in acute ischemic stroke with early endovascular therapy. Front Neurol. 2022;13:906377.
20. Zhang Z, Pu Y, Mi D, Liu L. Cerebral hemodynamic evaluation after cerebral recanalization therapy for acute ischemic stroke. Front Neurol. 2019;10:719.
21. Baracchini C, Farina F, Palmieri A, et al. Early hemodynamic predictors of good outcome and reperfusion injury after endovascular treatment. Neurology. 2019;92:e2774-83.
22. Kneihsl M, Niederkorn K, Deutschmann H, et al. Increased middle cerebral artery mean blood flow velocity index after stroke thrombectomy indicates increased risk for intracranial hemorrhage. J Neurointerv Surg. 2018;10:882-7.
23. Perren F, Kargiotis O, Pignat J-M, Pereira VM. Hemodynamic changes may indicate vessel wall injury after stent retrieval thrombectomy for acute stroke. Journal of Neuroimaging. 2018;28:412-5.
24. Wu S, Yuan R, Wang Y, et al. Early prediction of malignant brain edema after ischemic stroke. Stroke. 2018;49:2918-27.
25. Ng FC, Yassi N, Sharma G, et al. Cerebral edema in patients with large hemispheric infarct undergoing reperfusion treatment: a HERMES meta-analysis. Stroke. 2021;52:3450-8.
26. Tian B, Tian X, Shi Z, et al. Clinical and imaging indicators of hemorrhagic transformation in acute ischemic stroke after endovascular thrombectomy. Stroke. 2022;53:1674-81.
27. Silverman A, Kodali S, Sheth KN, Petersen NH. Hemodynamics and hemorrhagic transformation after endovascular therapy for ischemic stroke. Front Neurol. 2020;11:728.
28. Andrade JBC, Mohr JP, Lima FO, et al. The role of hemorrhagic transformation in acute ischemic stroke upon clinical complications and outcomes. J Stroke Cerebrovasc Dis. 2020;29:104898.
29. Bourcier R, Saleme S, Labreuche J, et al. More than three passes of stent retriever is an independent predictor of parenchymal hematoma in acute ischemic stroke. J Neurointerv Surg. 2019;11:625-9.
30. Mistry EA, Mistry AM, Nakawah MO, et al. Systolic blood pressure within 24 hours after thrombectomy for acute ischemic stroke correlates with outcome. J Am Heart Assoc. 2017;6:e006167.
31. Kim TJ, Park HK, Kim JM, et al. Blood pressure variability and hemorrhagic transformation in patients with successful recanalization after endovascular recanalization therapy: a retrospective observational study. Ann Neurol. 2019;85:574-81.
32. Fransen PS, Beumer D, Berkhemer OA, et al. MR CLEAN, a multicenter randomized clinical trial of endovascular treatment for acute ischemic stroke in the Netherlands: study protocol for a randomized controlled trial. Trials. 2014;15:343.
33. Demchuk AM, Goyal M, Menon BK, et al. Endovascular treatment for Small Core and Anterior circulation Proximal occlusion with Emphasis on minimizing CT to recanalization times (ESCAPE) trial: methodology. Int J Stroke. 2015;10:429-38.
34. Molina CA, Chamorro A, Rovira A, et al. REVASCAT: a randomized trial of revascularization with SOLITAIRE FR device vs. best medical therapy in the treatment of acute stroke due to anterior circulation large vessel occlusion presenting within eight-hours of symptom onset. Int J Stroke. 2015;10:619-26.

DOI: 10.4274/forbes.galenos.2026.44342
Forbes J Med 2026;7:69-74

Original Article

Clinical Characteristics, Indications, and Outcomes of the Low-Dose ACTH Stimulation Test in a Pediatric Endocrinology Center

Berna EROĞLU FİLİBELİ[1], Gülümay VURAL TOPAKTAŞ[1], Bumin N. DÜNDAR[2]

[1]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Pediatric Endocrinology, İzmir, Türkiye
[2]İzmir Katip Çelebi University Faculty of Medicine, Department of Pediatric Endocrinology, İzmir, Türkiye

Cite as: Eroğlu Filibeli B, Topaktaş Vural G, Dündar BN. Clinical characteristics, indications, and outcomes of the low-dose ACTH stimulation test in a pediatric endocrinology center. Forbes J Med. 2026;7:69-74

ABSTRACT

Objective: The low-dose adrenocorticotropic hormone stimulation test (LDAT) is widely used to evaluate the hypothalamic-pituitary-adrenal axis in pediatric endocrinology, particularly in suspected central adrenal insufficiency. This study aimed to retrospectively evaluate the clinical characteristics, indications, and outcomes of LDAT in a tertiary pediatric endocrinology center.

Methods: This retrospective study included 120 pediatric patients who underwent LDAT. A peak serum cortisol level ≥18 µg/dL at 30 minutes was considered an adequate adrenal response. Demographic, clinical, and laboratory data were analyzed. Receiver operating characteristic (ROC) analysis was performed to assess the diagnostic performance of baseline cortisol levels.

Results: A total of 120 patients (61 females, 59 males) were included. The median age was 11.5 years (interquartile range: 5.6); 55.8% were pubertal. The most common indication for testing was suspected central hypothyroidism (30.8%), followed by short stature or growth hormone deficiency (27.5%), and intracranial lesions or malformations (21.7%). An insufficient cortisol response was observed in 24.2% of patients. No significant differences were found between patients with adequate and insufficient responses regarding age, sex, anthropometric characteristics, or test indications ($p>0.05$). Baseline cortisol levels were significantly higher in patients with adequate responses (9.02±3.73 vs. 5.33±3.94 µg/dL, $p<0.001$). ROC analysis showed moderate diagnostic performance (area under the curve =0.75): a cut-off value of 5.05 µg/dL yielded 86.8% sensitivity and 55.2% specificity. None of the patients with baseline cortisol >13 µg/dL had an insufficient response.

Conclusion: LDAT remains a practical tool for evaluating adrenal function in children. Baseline cortisol may help identify low-risk patients, but dynamic testing remains essential for definitive diagnosis.

Keywords: Low-dose ACTH test, central adrenal insufficiency, pediatric endocrinology, cortisol, HPA axis

Received: 17.05.2026
Accepted: 28.06.2026
Epub: 29.06.2026
Publication Date: 09.07.2026

Corresponding Author:
Berna EROĞLU FİLİBELİ
University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Pediatric Endocrinology, İzmir, Türkiye
bernafilibeli@gmail.com
ORCID ID: 0000-0002-2696-0195

INTRODUCTION

Accurate assessment of the hypothalamic-pituitary-adrenal (HPA) axis is essential for the diagnosis of central adrenal insufficiency (CAI). Identifying patients with impaired adrenal function is crucial to prevent life-threatening complications, such as adrenal crisis, while avoiding overdiagnosis in individuals with a normally functioning axis is equally important. The insulin tolerance test (ITT) has traditionally been considered the gold standard for evaluating the HPA axis; however, its routine use in pediatric practice is limited by the risks and discomfort associated with induced hypoglycemia.[1] In addition, ITT is contraindicated in patients with cardiovascular disease or with a history of seizures.[2] Therefore, safer alternative dynamic tests are frequently used in children. The human corticotropin-releasing hormone test has been proposed as an acceptable method for assessing the HPA axis because it stimulates the axis at a higher anatomical level and induces endogenous adrenocorticotropic hormone (ACTH) secretion. In contrast, synthetic ACTH stimulation tests evaluate adrenal responsiveness to exogenous ACTH. Among these, the low-dose ACTH stimulation test (LDAT) is considered to provide a more physiological adrenal stimulus than the standard-dose test and has been suggested to be more sensitive for detecting CAI.[2]

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

Although several studies have proposed different threshold values for morning cortisol levels that may indicate a functionally intact HPA axis, there is no clear consensus.[3-5] In pediatric populations, LDAT is frequently required to assess adrenal function. The aim of this study was to retrospectively evaluate the demographic characteristics, indications for testing, and results of patients who underwent LDAT in a tertiary pediatric endocrinology center.

METHODS

This retrospective study included pediatric patients who underwent an LDAT at a tertiary pediatric endocrinology center between November 2023 and January 2026. The study was approved by the University of Health Sciences Türkiye, İzmir City Hospital Ethics Committee (approval no: 2026/225, date: 08.04.2026). Written informed consent was obtained from the parents or legal guardians of all participants. All procedures were conducted in accordance with the Declaration of Helsinki.

Demographic data, clinical characteristics, indications for testing, and laboratory findings were obtained from medical records. Weight, height, and body mass index standard deviation scores were calculated according to age- and sex-specific national reference data.[6]

Basal serum cortisol and ACTH measurements were obtained from routine blood samples collected during outpatient follow-up visits and analyzed in the hospital's central laboratory according to standard clinical practice. In contrast, baseline (0-minute) ACTH and cortisol samples were collected under standardized test conditions in a dedicated testing room by experienced healthcare personnel trained in endocrine stimulation testing, as part of the LDAT protocol. All blood samples during the LDAT procedure were collected at protocol-defined time points to ensure procedural consistency and the reliability of test results. Serum ACTH and cortisol concentrations were measured by electrochemiluminescence immunoassay on a Cobas® e801 analyzer (Roche Diagnostics, Mannheim, Germany).

The LDAT was performed according to standard protocols. All tests were conducted in the morning hours (typically between 08:00 and 09:00). A 1 µg dose of synthetic ACTH (Synacthen) was administered intravenously. To prepare the 1 µg dose, a 250 µg Synacthen ampoule was diluted to a final volume of 250 mL with normal saline, and 1 mL of this solution was used for testing. Serum cortisol levels were measured at baseline (0 minutes) and at 30 minutes after ACTH administration.[2] Although cortisol levels at 0 and 30 minutes were available for all patients undergoing LDAT, basal cortisol and ACTH measurements prior to testing were not obtained for all patients because of the retrospective nature of the study. A peak serum cortisol level ≥18 µg/dL at 30 minutes was accepted as an adequate adrenal response, whereas values <18 µg/dL were considered indicative of an insufficient cortisol response.[7] Patients were classified into adequate and insufficient response groups.

Statistical Analysis

Statistical analyses were performed using SPSS software (version 24.0; IBM Corp., Armonk, NY, USA). The distribution of continuous variables was assessed using the Shapiro-Wilk test. Continuous variables were tested for normality using the Shapiro-Wilk test and presented as mean ± standard deviation or median [interquartile range (IQR)], as appropriate.

Comparisons between groups (adequate vs. insufficient cortisol response) were performed using the Mann-Whitney U test for continuous variables and the chi-square test or Fisher's exact test for categorical variables, as appropriate.

Receiver operating characteristic (ROC) curve analysis was conducted to evaluate the diagnostic performance of cortisol levels measured at 0 minutes. The area under the curve (AUC) and the 95% confidence intervals (CIs) were calculated. The optimal cut-off value was determined based on the best balance between sensitivity and specificity. A p value <0.05 was considered statistically significant.

RESULTS

A total of 120 patients (61 females, 59 males) underwent the LDAT. The median age was 11.5 years (IQR: 5.6); 55.8% of the cohort were in the pubertal stage. The anthropometric characteristics and baseline ACTH and cortisol levels are summarized in Table 1.

The most common indications for testing were thyroid function abnormalities suggestive of central hypothyroidism (30.8%), followed by short stature or growth hormone deficiency (27.5%), intracranial mass lesions or malformations (21.7%), and other anterior pituitary hormone abnormalities (6.7%) (Figure 1). Less frequent indications included sepsis, prolonged steroid use, and hypoglycemia. Additional anterior pituitary hormone deficiencies were present in 30.3% of the patients. Among patients with additional anterior pituitary hormone deficiencies, growth hormone deficiency was the most frequently observed (57.9%).

An insufficient cortisol response was observed in 24.2% of the patients (n=29). No statistically significant differences were observed between the adequate and insufficient response groups with respect to age, sex, anthropometric characteristics, pubertal status, indications for testing,

or presence of additional anterior pituitary hormone deficiencies (all p>0.05) (Table 2).

A total of 37 patients underwent testing for suspected central hypothyroidism, 10 of whom exhibited an insufficient cortisol response. Among these patients, five were obese, two were malnourished, and three had a history of antipsychotic or antiepileptic drug use that was associated with thyroid function test abnormalities.

The mean serum cortisol level at 0 minutes was significantly higher in the adequate response group compared to the insufficient response group (9.02±3.73 µg/dL vs. 5.33±3.94 µg/dL, p<0.001).

ROC analysis demonstrated that cortisol levels at 0 minutes had a moderate ability to discriminate between adequate and insufficient cortisol responses (AUC =0.75, 95% CI: 0.65-0.85, p<0.001) (Figure 2). A cut-off value of 5.05 µg/dL yielded a sensitivity of 86.8% and a specificity of 55.2% for predicting an adequate cortisol response. Notably, none of the patients with a baseline cortisol level above 13 µg/dL exhibited an insufficient cortisol response during LDAT.

Table 1. Baseline characteristics of the study population

Variable	n	Median (IQR)
Age (years)	120	11.51 (5.59)
Weight (kg)	120	45.0 (47.15)
Weight SDS	120	0.25 (2.60)
Height (cm)	120	147.0 (36.15)
Height SDS	120	-0.25 (1.35)
BMI (kg/m²)	120	20.22 (9.51)
BMI SDS	120	0.57 (2.30)
Basal ACTH (pg/mL)	73	17.8 (13.97)
Basal cortisol (µg/dL)	89	6.9 (4.90)

IQR: Interquartile range, BMI: Body mass index, SDS: Standard deviation scores, ACTH: Adrenocorticotropic hormone

DISCUSSION

In this study, we evaluated the clinical characteristics, indications, and outcomes of LDAT in a pediatric population at a tertiary endocrinology center. In our cohort, the most frequent indication for testing was to rule out concomitant CAI in patients diagnosed with central hypothyroidism prior to the initiation of therapy. Although this is not a typical indication for LDAT, it may be explained by important clinical considerations. In patients with suspected central hypothyroidism, adrenal function should be evaluated before initiating levothyroxine replacement, as untreated concomitant adrenal insufficiency may precipitate life-threatening adrenal crisis following thyroid hormone therapy.[8] This issue is particularly relevant in pediatric endocrinology and may account for the high frequency with which LDAT was performed for this indication in our cohort. On the other hand, central hypothyroidism is a relatively rare endocrine disorder.[9] Furthermore, biochemical findings suggestive of central hypothyroidism—particularly low free T4 levels in the presence of normal or low TSH—may be influenced by multiple factors, including assay variability, laboratory interference, medication use, and timing of sampling, potentially leading to misinterpretation of thyroid function tests.[10-12] In our clinical practice, thyroid function tests are often repeated two or three times to confirm the persistence of abnormalities before making a definitive diagnosis. In patients with persistent findings suggestive of central hypothyroidism, LDAT was performed

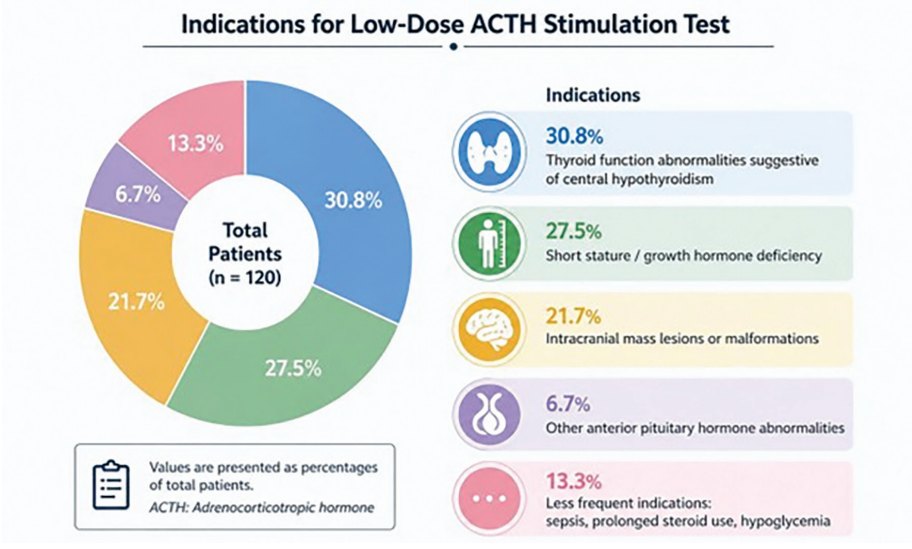


Figure 1. Distribution of indications for the low-dose ACTH stimulation test

to exclude concomitant CAI. In addition, free T4 levels are measured in our laboratory using immunoassay methods, which are known to be susceptible to analytical interference. In contrast, liquid chromatography-tandem mass spectrometry provides more accurate and specific measurements of thyroid hormones, particularly in the presence of assay-related limitations.[13] Taken together, these factors may contribute to the high frequency of suspected central hypothyroidism as an indication for LDAT in our cohort. The use of more accurate laboratory methods may reduce unnecessary testing in these patients, thereby minimizing both healthcare costs and the associated psychological burden on families.

Other frequent indications for testing included short stature/growth hormone deficiency, intracranial mass lesions or malformations, and other anterior pituitary hormone abnormalities, which further support the importance of evaluating adrenal function in these patients. Our findings demonstrate that LDAT is widely used across diverse clinical indications and that approximately one-quarter of patients exhibit an insufficient cortisol response. Consistent with previous reports, we found no significant differences between groups in terms of age, sex, or anthropometric parameters.[9,14] This suggests that clinical characteristics alone are insufficient to predict adrenal insufficiency, reinforcing the need for biochemical evaluation. Similarly, the presence of additional anterior pituitary hormone deficiencies did not differ significantly between groups, indicating that adrenal insufficiency cannot be reliably inferred, even in patients with multiple pituitary abnormalities.

One of the key findings of our study is that cortisol levels measured at 0 minutes were moderately predictive of adrenal sufficiency, with an AUC of 0.75. A cut-off value of 5.05 µg/dL provided high sensitivity (86.8%) but moderate specificity. This suggests that baseline cortisol measurements obtained at the time of testing may have clinical utility in identifying patients at low-risk of adrenal insufficiency, potentially reducing the need for dynamic testing in selected cases. However, the relatively low

Table 2. Comparison of clinical, anthropometric, and hormonal characteristics between patients with adequate and insufficient cortisol response

Variable	Adequate response (n=91)	Insufficient response (n=29)	p value
Age (years)	11.6 (5.34)	10.64 (5.45)	0.38[a]
Sex (female/male)	50/41	11/18	0.11[b]
Weight (kg)	45.24±28.39	45.58±34.1	0.46[b]
Weight SDS	0.5±1.87	0.31±1.65	0.90[b]
Height (cm)	147.8 (35.47)	134.0 (44)	0.66[a]
Height SDS	-0.30±1.18	-0.52±0.96	0.81[b]
BMI (kg/m²)	22.41±7.8	21.41±9.04	0.75[b]
BMI SDS	0.56±1.71	0.56±1.63	0.79[b]
Prepubertal/pubertal	33/58	9/20	0.93[b]
Etiology (%)			0.69[b]
-Suspected central hypothyroidism	29.7	34.5	
-Intracranial mass lesions or malformations	19.8	27.6	
-Short stature/growth hormone deficiency	29.7	20.7	
-Other anterior pituitary hormone abnormalities	5.5	3.4	
Additional pituitary hormone deficiency (%)	29.7	31.0	0.91[b]
Basal ACTH (n, 7.2-63.3 pg/mL) (n=73)	16.35 (13.36)	22.4 (8.4)	**0.03[b]**
Basal cortisol (µg/dL) (n=89)	6.57±3.78	6.01±3.37	0.34[a]
ACTH (0 min) (n, 7.2-63.3 pg/mL)	28.36±18.3	22.87±14.46	0.42[b]
Cortisol (0 min) (µg/dL)	9.02±3.73	5.33±3.94	**<0.001[b]**
Cortisol (30 min) (µg/dL)	20.1 (4.92)	13.9 (2.8)	**<0.001[a]**

Values are presented as mean ± standard deviation or median (interquartile range).

[a]Mann-Whitney U test

[b]Chi-square test or Student's t-test, as appropriate

BMI: Body mass index, SDS: Standard deviation scores, ACTH: Adrenocorticotropic hormone

specificity indicates that such an approach should be used cautiously and in conjunction with clinical judgment. These findings are in line with previous studies reporting considerable variability in the diagnostic performance of basal cortisol levels, with no universally accepted cut-off value.[2,4,5] Although higher morning cortisol levels (e.g., >13 µg/dL) are generally considered indicative of an intact HPA axis, values within the intermediate range typically require further evaluation with dynamic stimulation tests.[2,3] Notably, none of the patients in our cohort with a baseline cortisol level above 13 µg/dL exhibited an insufficient cortisol response during LDAT, further supporting the clinical utility of this threshold for excluding adrenal insufficiency.

Although basal ACTH levels were significantly higher in the insufficient response group, levels remained within the reference range in both groups. Similarly, basal cortisol levels obtained before testing did not differ significantly between groups, whereas concentrations measured at 0 minutes during LDAT were significantly higher in the adequate response group. These findings may be explained, at least in part, by differences in sampling conditions. Routine outpatient ACTH and cortisol measurements are susceptible to preanalytical factors, particularly variations in sampling time and collection procedures. In routine clinical practice, blood samples may be obtained at different times of the day and under varying conditions with respect to fasting status, recent physical activity, and psychological stress, all of which may influence cortisol concentrations. In contrast, ACTH and cortisol samples obtained during LDAT were collected under standardized conditions according to a predefined protocol. Notably, ACTH concentrations measured at the start of the stimulation test did not differ significantly between groups. In addition, the limited number of available basal ACTH and cortisol measurements may have reduced the statistical power of these comparisons. Therefore, the observed differences in routine basal hormone measurements should be interpreted cautiously and are unlikely to reflect clinically meaningful differences in HPA axis function. These findings further emphasize the importance of standardized sampling conditions when interpreting ACTH and cortisol levels.

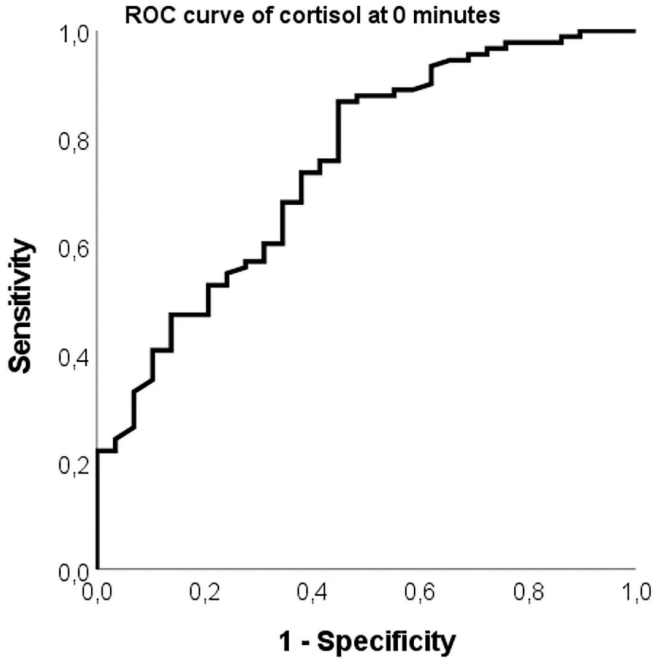


Figure 2. Receiver operating characteristic (ROC) curve of cortisol levels at 0 minutes for predicting adequate adrenal response

The use of a peak cortisol cut-off of 18 µg/dL is consistent with widely accepted clinical thresholds; however, assay-specific variability may influence optimal cut-off values. Recent studies have suggested that lower thresholds may be appropriate when using modern immunoassays or mass spectrometry-based methods.[2,15,16] Therefore, caution is warranted when generalizing cut-off values across different laboratory settings.

Study Limitations

This study has several limitations. First, its retrospective design may have introduced selection bias, reduced control over preanalytical conditions, and limited the availability of certain laboratory parameters, including basal ACTH and cortisol levels. In addition, the single-center nature of the study may limit the generalizability of the findings to other populations and clinical settings. Second, we did not perform confirmatory testing with ITT or other dynamic tests, which limits the ability to directly compare diagnostic performance. Third, assay-specific variability was not evaluated, and differences in cortisol measurement methods across laboratories may influence the applicability of the proposed cut-off values. Despite these limitations, our study provides valuable real-world data on the use of LDAT in a pediatric population and highlights the potential role of baseline cortisol levels obtained during testing.

CONCLUSION

LDAT remains a practical and widely used method for evaluating adrenal function in children. Baseline cortisol levels measured at the time of testing may provide useful clinical information for identifying patients at low-risk of adrenal insufficiency; however, dynamic testing remains essential for definitive diagnosis. Further prospective studies are needed to refine diagnostic thresholds and optimize testing strategies in pediatric populations.

Ethics

Ethics Committee Approval: The study was approved by the University of Health Sciences Türkiye, İzmir City Hospital Ethics Committee (approval no: 2026/225, date: 08.04.2026).

Informed Consent: Written informed consent was obtained from the parents or legal guardians of all participants.

Acknowledgment

The authors thank all healthcare staff involved in the clinical care of the patients included in this study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.E.F., G.T.V., Concept: B.E.F., Design: B.E.F., B.N.D., Data Collection or Processing: G.T.V., Analysis or Interpretation: B.E.F., G.T.V., B.N.D., Literature Search: B.E.F., G.T.V., B.N.D., Writing: B.E.F., G.T.V., B.N.D.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Bornstein SR, Allolio B, Arlt W, et al. Diagnosis and treatment of primary adrenal insufficiency: an Endocrine Society Clinical Practice guideline. J Clin Endocrinol Metab. 2016;101:364-89.
2. Yıldız M, Yıldırım R, Baş F. Central adrenal insufficiency: etiology and diagnostic approach. J Clin Res Pediatr Endocrinol. 2025;17:72-9.
3. Kazlauskaite R, Maghnie M. Pitfalls in the diagnosis of central adrenal insufficiency in children. Endocr Dev. 2010;17:96-107.
4. Dullaart RP, Pasterkamp SH, Beentjes JA, Sluiter WJ. Evaluation of adrenal function in patients with hypothalamic and pituitary disorders: comparison of serum cortisol, urinary free cortisol and the human-corticotrophin releasing hormone test with the insulin tolerance test. Clin Endocrinol (Oxf). 1999;50:465-71.
5. Maguire AM, Biesheuvel CJ, Ambler GR, Moore B, McLean M, Cowell CT. Evaluation of adrenal function using the human corticotrophin-releasing hormone test, low dose Synacthen test and 9am cortisol level in children and adolescents with central adrenal insufficiency. Clin Endocrinol (Oxf). 2008;68:683-91.
6. Neyzi O, Bundak R, Gökçay G, et al. Reference values for weight, height, head circumference, and body mass index in Turkish children. J Clin Res Pediatr Endocrinol. 2015;7:280-93.
7. Dickstein G, Shechner C, Nicholson WE, et al. Adrenocorticotropin stimulation test: effects of basal cortisol level, time of day, and suggested new sensitive low dose test. J Clin Endocrinol Metab. 1991;72:773-8.
8. Persani L, Brabant G, Dattani M, et al. 2018 European Thyroid Association (ETA) guidelines on the diagnosis and management of central hypothyroidism. Eur Thyroid J. 2018;7:225-37.
9. Crowley RK, Argese N, Tomlinson JW, Stewart PM. Central hypoadrenalism. J Clin Endocrinol Metab. 2014;99:4027-36.
10. Persani L. Clinical review: central hypothyroidism: pathogenic, diagnostic, and therapeutic challenges. J Clin Endocrinol Metab. 2012;97:3068-78.
11. Favresse J, Burlacu MC, Maiter D, Gruson D. Interferences with thyroid function immunoassays: clinical implications and detection algorithm. Endocr Rev. 2018;39:830-50.
12. Fliers E, Bianco AC, Langouche L, Boelen A. Thyroid function in critically ill patients. Lancet Diabetes Endocrinol. 2015;3:816-25.
13. Soldin OP, Soldin SJ. Thyroid hormone testing by tandem mass spectrometry. Clin Biochem. 2011;44:89-94.
14. Ospina NS, Al Nofal A, Bancos I, et al. ACTH stimulation tests for the diagnosis of adrenal insufficiency: systematic review and meta-analysis. J Clin Endocrinol Metab. 2016;101:427-34.
15. Dickstein G, Saiegh L. Low-dose and high-dose adrenocorticotropin testing: indications and shortcomings. Curr Opin Endocrinol Diabetes Obes. 2008;15:244-9.
16. El-Farhan N, Rees DA, Evans C. Measuring cortisol in serum, urine and saliva - are our assays good enough? Ann Clin Biochem. 2017;54:308-22.

DOI: 10.4274/forbes.galenos.2026.60437
Forbes J Med 2026;7:75-83

Original Article

Impact of Preprocedural Ultrasound-Guided Marking on Spinal Anesthesia Success in Elderly Patients Undergoing Lower Extremity Orthopedic Surgery: A Retrospective Study

Hazal KOCA, Hale AKSU ERDOST, Semih KÜÇÜKGÜÇLÜ

Dokuz Eylül University Faculty of Medicine, Department of Anesthesiology and Reanimation, İzmir, Türkiye

Cite as: Koca H, Aksu Erdost H, Küçükgüçlü S. Impact of preprocedural ultrasound-guided marking on spinal anesthesia success in elderly patients undergoing lower extremity orthopedic surgery: a retrospective study. Forbes J Med. 2026;7:75-83

ABSTRACT

Objective: Anatomical and physiological changes in geriatric patients may make spinal anesthesia technically challenging. The traditional manual palpation technique is a blind approach, and anatomical variations may reduce block success and increase complications. Ultrasonography can provide guidance for spinal anesthesia. This study investigated the effect of preprocedural ultrasound-guided marking on block success in patients aged ≥65 years undergoing lower extremity orthopedic surgery.

Methods: This retrospective observational study included 180 patients aged ≥65 years who underwent lower extremity orthopedic surgery under spinal anesthesia at Dokuz Eylül University Faculty of Medicine Hospital. Of these, 88 patients were in the ultrasound-guided marking group and 92 in the manual palpation group. The primary outcome was first-attempt success of spinal anesthesia. Secondary outcomes included puncture attempts, needle redirections, time to return of cerebrospinal fluid, and spinal needle placement time.

Results: Demographic characteristics were similar between the groups. The numbers of spinal punctures and needle redirections were significantly lower in the ultrasound group. In addition, spinal needle placement time was significantly shorter in the ultrasound group ($p<0.001$). Successful spinal anesthesia with a single puncture was achieved in 73.9% of patients in the ultrasound group and in 44.6% of patients in the manual palpation group ($p<0.001$).

Conclusion: Preprocedural ultrasound-guided marking was associated with higher first-attempt success, fewer puncture attempts, fewer needle redirections, and shorter needle-placement time in elderly patients undergoing spinal anesthesia for lower extremity orthopedic surgery. Given the retrospective non-randomized design, these findings should be interpreted as an association rather than a definitive causal effect.

Keywords: Spinal anesthesia, neuraxial block, ultrasonography, elderly, orthopedic surgery

Received: 16.03.2026
Accepted: 09.07.2026
Epub: 13.07.2026
Publication Date: 13.08.2026

Corresponding Author:
Hazal KOCA
Dokuz Eylül University Faculty of Medicine, Department of Anesthesiology and Reanimation, İzmir, Türkiye
hazalkucukali67@gmail.com
ORCID ID: 0000-0003-2264-108X

INTRODUCTION

Technological advancements have increased average lifespans worldwide, resulting in a greater need for anesthesia among geriatric patients. Procedures such as hip replacement, knee replacement, and hip fracture repair are commonly performed in this population. The decreased physiological adaptive capacity and the presence of concomitant systemic diseases in geriatric patients increase the risk of perioperative complications. Regional anesthesia has been shown to reduce complications, mortality, and morbidity among geriatric patients. However, there is still no clear consensus regarding the safest anesthesia method for these patients.[1-3]

Regional anesthesia is frequently preferred for lower extremity surgeries in geriatric patients. It has the potential to reduce cardiovascular and pulmonary complications compared with general anesthesia. Spinal anesthesia is one of the most commonly used regional anesthesia techniques for lower extremity surgeries. However, this technique may present certain challenges in elderly patients. Age-related changes, such as increased spinal curvature (dorsal kyphosis), vertebral collapse,

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

intervertebral disc dehydration, osteophyte formation, and calcification of the supraspinous and interspinous ligaments and the ligamentum flavum, make midline approaches more difficult. In addition, decreased epidural resistance, reduced epidural fat, and foraminal narrowing increase epidural compliance, which may lead to higher block levels with the same volume of local anesthetic. Sclerotic changes and calcifications in the intervertebral foramina can also narrow the foraminal space, and the ligamentum flavum may calcify. These anatomical changes, together with difficulties in positioning elderly patients, make central spinal blocks particularly challenging.[3,4] The success of spinal anesthesia depends on the experience of the anesthesia provider, the anatomical characteristics of the patient, and the techniques used. Therefore, accurate identification of anatomical landmarks is essential for successful spinal anesthesia.

In the traditional manual palpation technique, the performance of the spinal block is based on palpation of the patient's anatomical landmarks. This blind approach can be affected by anatomical variability and may result in decreased block success and increased complications. In elderly patients, anatomical landmarks may be less prominent, thereby making identification of the correct puncture site more difficult. Ultrasound enables visualization of the spinal structures, needle insertion point, and surrounding anatomy. It allows identification of the spinal midline, intervertebral levels, optimal needle insertion points, the appropriate needle trajectory, as well as estimation of the depth of the epidural space.[5,6] Ultrasound is a safe, validated, noninvasive, and relatively easy-to-learn method that has been used as a more reliable tool for identifying intervertebral levels. Previous studies have shown that pre-procedural ultrasound scanning can shorten the duration of spinal anesthesia procedures, reduce the number of punctures, and improve first-attempt success rates, particularly in patients with poorly palpable surface landmarks. Additionally, ultrasound-guided marking may help reduce complications associated with spinal anesthesia.[7-11]

The aim of this study was to investigate the effect of ultrasonographic marking performed before spinal anesthesia on the success of spinal anesthesia in patients aged 65 years and older undergoing lower extremity surgery. We evaluated the effect of ultrasound on spinal needle placement time, first-attempt success rate, number of punctures, and early complication rates by comparing patients whose neuraxial anesthesia sites were marked with ultrasound versus those marked by manual palpation.

METHODS

Study Design and Ethics Approval

This retrospective study was conducted at Dokuz Eylül University Faculty of Medicine Hospital after obtaining approval from the Non-Interventional Ethics Committee (approval number: 2018/17-36, date: 18.06.2019). The study included patients aged 65 years and older who underwent lower extremity orthopedic surgery under spinal anesthesia. In this retrospective study, the success of spinal anesthesia was compared between two groups: patients who underwent ultrasound-guided marking before spinal puncture and those in whom the puncture site was determined by manual palpation. Patient allocation to the ultrasound or manual palpation groups was determined by routine clinical practice rather than by randomization. Because the study was retrospective, an a priori sample size calculation was not performed; however, all eligible patients during the study period were included in the analysis. The study was conducted in accordance with the STROBE guidelines for observational studies.

Data Collection

Pre-anesthetic evaluation data (physical examination findings, informed consent forms and anesthesia records) were retrieved from patient files and the hospital information system. The following parameters were used to assess spinal anesthesia success:

- Number of spinal punctures (defined as each needle entry through the skin, regardless of whether it occurred at the same or a different intervertebral level).
- Number of needle redirections (defined as repositioning of the needle without complete withdrawal from the skin).
- Time elapsed from the initial needle insertion to the appearance of cerebrospinal fluid (CSF).
- Spinal needle-placement time (defined as the time from the first needle insertion to needle withdrawal following drug administration into the subarachnoid space). This variable reflects the duration of the needle-placement phase and does not include preprocedural steps such as ultrasound scanning or sterile preparation.

The primary outcome was first-attempt success of spinal anesthesia. Secondary outcomes included the number of puncture attempts, the number of needle redirections, time to CSF return, and spinal needle placement time.

Demographic data, including height, weight, body mass index (BMI), surgical type, and American Society of Anesthesiologists (ASA) classification, were recorded. Additional procedural data, including spinal needle

gauge, number of spinal punctures, number of needle redirections, time from needle insertion to CSF flow, and spinal needle placement time, were also recorded. Early complications such as paresthesia during needle insertion and the presence of blood in the needle during puncture were documented. Anesthesiologists' experience was defined as the number of years since the start of anesthesiology residency training. Spinal anesthesia procedures were performed by anesthesia residents with at least two years' experience, under the supervision of attending anesthesiologists.

Ultrasound-Guided Marking Technique

The preprocedural ultrasound examination was performed by the same anesthesiologist, who had six years' experience in neuraxial ultrasonography. Ultrasound imaging was performed using a LOGIQ-E device (GE Medical Systems, China) with a 3.5-MHz convex probe.

Patients were placed in the sitting position, and the lumbar spine was scanned from the sacrum toward the Tuffier line in longitudinal, paramedian, and sagittal views. The L4-L5 interspinous space was identified based on the ultrasound appearance of the spinous processes. The optimal needle insertion point was marked on the skin using a surgical marker. The transverse interlaminar ultrasound view was used to identify the midline of the vertebral column.

After ultrasound-guided marking, patients were transferred to the operating room where routine monitoring was performed according to ASA standards. Sedation with midazolam and/or fentanyl was administered during the procedure according to routine clinical practice.

Following skin antisepsis and sterile draping, local anesthesia was applied to the skin and subcutaneous tissue at the marked point. A 90-mm Quincke spinal needle (Braun, Germany) of either 25-G or 27-G was used. After free flow of CSF was observed, 12-14 mg of hyperbaric bupivacaine with 25 μg of fentanyl was administered intrathecally.

Manual Palpation Technique

In the manual palpation group, patients were prepared in the operating room with routine ASA-standard monitoring, establishment of intravenous access, and fluid administration. Patients were placed in the sitting position and the needle insertion site was determined by palpation of the iliac crests and spinous processes. The interspinous space corresponding to the Tuffier line was considered the L4-L5 level and was used as a guide for the spinal puncture.

Following skin antisepsis and sterile draping, local anesthesia was administered to the skin and subcutaneous tissue at the identified puncture site. As in the ultrasound group, a 25-G or 27-G, 90-mm Quincke spinal needle (Braun, Germany) was used. After confirmation of free flow of CSF, 12-14 mg of hyperbaric bupivacaine with 25 μg of fentanyl was administered intrathecally.

Exclusion Criteria

Patients younger than 65 years, those undergoing emergency surgery, non-Turkish-speaking patients, patients with contraindications to regional anesthesia, patients with a history of vertebral surgery, and patients requiring conversion to general anesthesia were excluded from the study.

In addition, patients with incomplete or missing data regarding spinal anesthesia procedure parameters were excluded.

Patients with non-standard sedation dosing or incomplete sedation records were also excluded to ensure consistency of the dataset.

Statistical Analysis

Statistical analysis was performed using SPSS version 23.0. The normality of numerical variables was assessed using the Kolmogorov-Smirnov test. For normally distributed variables, the independent-samples t-test was used, whereas the Mann-Whitney U test was applied to non-normally distributed variables. Categorical variables were compared between groups using the chi-square test. Independent t-tests were used for demographic variables, spinal block-related parameters, and procedure times, whereas first-attempt success was analyzed using the chi-square test. A p value <0.05 was considered statistically significant.

RESULTS

The study included 180 patients: 88 in the ultrasound-guided marking group and 92 in the manual palpation group.

Baseline demographic and clinical characteristics are presented in Table 1. No statistically significant differences were observed between the two groups with respect to age, weight, height, BMI, sex distribution, ASA classification, or type of surgery ($p>0.05$ for all comparisons).

The distributions of needle types and anesthesiologists' experience levels across the groups are presented in Table 2B. No statistically significant differences were observed between the groups for these variables ($p>0.05$).

Procedural outcomes are summarized in Table 2A. The numbers of puncture attempts and needle redirections were significantly lower in the ultrasound group compared with the manual palpation group ($p<0.001$ for both). Time to CSF flow and spinal needle placement time were also

Table 1. Baseline demographic and clinical characteristics of the study groups			
Variable	**Ultrasound group (n=88)**	**Manual palpation group (n=92)**	**p value**
Age (years), mean ± SD	73.56±7.44	73.33±7.29	0.834
Weight (kg), mean ± SD	77.78±12.33	75.56±13.22	0.246
Height (cm), mean ± SD	166.00±9.55	166.11±11.03	0.938
BMI (kg/m²), mean ± SD	28.46±5.50	27.46±4.90	0.199
BMI ≥30 kg/m², n (%)	32 (36.4%)	24 (26.1%)	0.137
BMI <30 kg/m², n (%)	56 (63.6%)	68 (73.9%)	
Male sex, n (%)	28 (31.8%)	31 (33.7%)	0.789
Female sex, n (%)	60 (68.2%)	61 (66.3%)	
ASA class, n (%)			0.940
ASA I	3 (3.4%)	4 (4.3%)	
ASA II	64 (72.7%)	67 (72.8%)	
ASA III	21 (23.9%)	21 (22.8%)	
Surgery type, n (%)			0.768
Hip/femur	33 (37.5%)	31 (33.7%)	
Knee/tibia/fibula	54 (61.4%)	59 (64.1%)	
Other	1 (1.1%)	2 (2.2%)	
Palpable spinous process, n (%)	54 (61.4%)	72 (78.3%)	**0.013***
Non-palpable spinous process, n (%)	34 (38.6%)	20 (21.7%)	

Data are presented as mean ± SD or number (percentage)
BMI: Body mass index, SD: Standard deviation, ASA: American Society of Anesthesiologists
*p<0.05 indicates statistical significance

Table 2A. Comparison of procedural outcomes			
Variable	**Ultrasound group (n=88)**	**Manual palpation group (n=92)**	**p value**
First-attempt success, n (%)	65 (73.9%)	41 (44.6%)	**<0.001***
Number of puncture attempts, mean ± SD	1.26±0.44	1.64±0.63	**<0.001***
Number of needle redirections, mean ± SD	2.11±1.62	4.36±2.59	**<0.001***
Time to CSF flow (s), mean ± SD	61.25±29.71	104.94±69.83	**<0.001***
Spinal needle-placement time (s), mean ± SD	95.06±36.73	144.80±77.16	**<0.001***

Data are presented as mean ± SD or number (percentage)
CSF: Cerebrospinal fluid, SD: Standard deviation
*p<0.05 indicates statistical significance

Table 2B. Technical characteristics of the procedure			
Variable	**Ultrasound group (n=88)**	**Manual palpation group (n=92)**	**p value**
Needle type, n (%)			0.498
25 G Quincke	81 (92.0%)	87 (94.6%)	
27 G Quincke	7 (8.0%)	5 (5.4%)	
Anesthesiologist experience (years), mean ± SD	2.50±0.23	2.46±0.26	0.247

Data are presented as mean ± SD or number (percentage)
SD: Standard deviation

significantly shorter in the ultrasound group (p<0.001 for both). In addition, the first-attempt success rate was significantly higher in the ultrasound group than in the manual palpation group (73.9% vs. 44.6%, p<0.001).

Among patients who achieved successful spinal anesthesia without needle redirection in a single puncture attempt (n=20), 14 (70.0%) were in the ultrasound group and 6 (30.0%) were in the manual palpation group. However, when analyzed within groups, the first-attempt success rate remained significantly higher in the ultrasound group (73.9% vs. 44.6%).

In subgroup analysis stratified by spinous process palpability (Table 3), puncture attempts, needle redirections, time to CSF flow, and spinal needle-placement time were significantly lower in the ultrasound group among patients with palpable spinous processes and among those with non-palpable spinous processes (p<0.001 for all comparisons).

Similarly, in subgroup analysis according to BMI (Table 4), all procedural parameters, including number of puncture attempts, needle redirections, time to CSF flow, and spinal needle-placement time, were significantly lower in the ultrasound group in both BMI ≥30 kg/m^2 and BMI <30 kg/m^2 subgroups (p<0.001 for all comparisons).

Early complications and sedation requirements are summarized in Table 5. The incidence of paresthesia did not differ significantly between the groups (p>0.05). However, the presence of blood in the spinal needle was significantly higher in the manual palpation group compared with the ultrasound group (p<0.001). In addition, doses of both midazolam and fentanyl were significantly lower in the ultrasound group than in the manual palpation group (p<0.001 for both).

DISCUSSION

This study analyzed retrospective data from geriatric patients who underwent lower extremity surgery under spinal anesthesia. In this retrospective study, preprocedural ultrasound-guided marking was associated with fewer puncture attempts, fewer needle redirections, shorter time to CSF flow, and shorter spinal needle-placement time compared with the manual palpation technique. Additionally, the success rate of achieving spinal anesthesia on the first puncture attempt was higher in the ultrasound group.

The aim of our study was to investigate whether pre-procedural ultrasound scanning improves the first-attempt success rate of spinal anesthesia in geriatric patients and whether it reduces the time required for the procedure, the number of punctures, and the number of needle redirections compared with the conventional palpation technique. The primary outcome of the study was first-attempt success of spinal anesthesia, while secondary outcomes included puncture attempts, needle redirections, time to CSF return, and spinal needle placement time. In overweight or obese patients, palpation of the spinous processes and iliac crests is more difficult, which may

Table 3. Subgroup analysis according to spinous process palpability

Variable	Ultrasound group (n=88)	Manual palpation group (n=92)	p value
Palpable spinous process: ultrasound group, n=54; manual palpation group, n=72			
Number of puncture attempts, mean ± SD	1.14±0.35	1.52±0.62	<0.001*
Number of needle redirections, mean ± SD	1.75±1.60	3.88±2.53	<0.001*
Time to CSF flow (s), mean ± SD	51.62±26.81	92.29±49.64	<0.001*
Spinal needle-placement time (s), mean ± SD	82.03±29.45	126.84±56.70	<0.001*
Non-palpable spinous process: ultrasound group, n=34; manual palpation group, n=20			
Number of puncture attempts, mean ± SD	1.44±0.50	2.05±0.51	<0.001*
Number of needle redirections, mean ± SD	2.67±1.51	6.10±3.06	<0.001*
Time to CSF flow (s), mean ± SD	76.52±27.94	150.50±106.56	<0.001*
Spinal needle-placement time (s), mean ± SD	115.76±38.02	202.25±109.83	<0.001*

Data are presented as mean ± SD
CSF: Cerebrospinal fluid, SD: Standard deviation
*p<0.05 indicates statistical significance

Table 4. Subgroup analysis according to body mass index			
Variable	**Ultrasound group (n=88)**	**Manual palpation group (n=92)**	**p value**
BMI ≥30 kg/m²: ultrasound group, n=32; manual palpation group, n=24			
Number of puncture attempts, mean ± SD	1.43±0.50	2.00±0.58	<0.001*
Number of needle redirections, mean ± SD	2.96±1.69	6.20±2.24	<0.001*
Time to CSF flow (s), mean ± SD	75.43±31.30	142.41±102.91	<0.001*
Spinal needle-placement time (s), mean ± SD	114.46±39.92	193.41±106.86	<0.001*
BMI <30 kg/m²: ultrasound group, n=56; manual palpation group, n=68			
Number of puncture attempts, mean ± SD	1.16±0.37	1.51±0.61	<0.001*
Number of needle redirections, mean ± SD	1.62±1.36	3.72±2.39	<0.001*
Time to CSF flow (s), mean ± SD	53.14±25.69	91.72±48.03	<0.001*
Spinal needle-placement time (s), mean ± SD	83.98±29.87	124.94±54.09	<0.001*

Data are presented as mean ± SD
BMI: body mass index, CSF: Cerebrospinal fluid, SD: Standard deviation
*p<0.05 indicates statistical significance

Table 5. Comparison of early complications and sedation requirements			
Variable	**Ultrasound group (n=88)**	**Manual palpation group (n=92)**	**p value**
Paresthesia, n (%)	6 (6.8%)	14 (15.2%)	0.073
Blood detected in needle, n (%)	2 (2.3%)	18 (19.6%)	<0.001*
Midazolam dosage (mg), mean ± SD	0.96±0.29	1.26±0.45	<0.001*
Fentanyl dosage (µg), mean ± SD	38.06±17.76	47.60±17.57	<0.001*

Data are presented as mean ± SD or number (percentage)
SD: Standard deviation
*p<0.05 indicates statistical significance

complicate identification of the appropriate puncture site using manual palpation.[9] The ability to palpate the spinous processes and to optimally position the patient to increase the intervertebral space is a critical factor for the success of spinal anesthesia. Reducing the number of punctures and needle redirections may improve the success rate of spinal anesthesia and reduce post-procedural complications.[5-7] In this study, palpation of the spinous processes was assessed as "palpable" or "non-palpable" without using a four-stage grading system. The retrospective nature of this study also precluded the use of randomization. Patient allocation was based on routine clinical practice rather than randomization, which introduced potential selection bias and residual confounding. This is particularly relevant because spinous process palpability differed between groups, a factor directly related to procedural difficulty. Although subgroup analyses were performed, residual confounding cannot be fully excluded. Another limitation of this study is the absence of a multivariable-adjusted analysis. Although baseline characteristics were broadly similar between groups, unmeasured confounders may have influenced the observed associations. Therefore, the results should be interpreted as an association rather than a definitive causal relationship.

Consistent with our findings, many randomized controlled trials have examined the use of ultrasound to facilitate spinal anesthesia in non-obstetric patients. Shaikh et al.[6] in their 2013 meta-analysis involving 1,334 patients across 14 studies, concluded that ultrasound imaging reduces the risk of failed or traumatic lumbar punctures, as well as the number of punctures and needle redirections. Consistent with these findings, the numbers of punctures and needle redirections were lower in our ultrasound group. Needle redirections were not counted as new puncture attempts.

Ultrasound-guided techniques significantly improve the likelihood of a successful puncture on the first attempt compared with conventional palpation techniques. Age-related anatomical changes, such as decreased vertebral height and narrowing of the intervertebral space, make needle placement more challenging. Tessler et al.[12] demonstrated that first-attempt success rates decrease with advancing age, reporting rates of 68.3%, 64.9%, and 61.9% in patients aged <50 years, 50-70 years, and >70

years, respectively. They identified advanced age as an independent risk factor affecting the success of spinal anesthesia. In our study, we included geriatric patients, a population known to be at higher risk of unsuccessful spinal anesthesia. Li et al.[9] compared manual palpation with pre-procedural ultrasound scanning in patients with a BMI ≥30 kg/m^2 scheduled for elective cesarean delivery. They reported that, among 80 patients, the ultrasound group had higher first-attempt success rates and required fewer punctures and needle redirections. Similarly, Khan et al.[8] examined 100 patients aged 18-60 years undergoing lower extremity orthopedic surgery with combined spinal-epidural anesthesia and found that the first-attempt success rate was significantly higher in the ultrasound group. Zeng et al.,[3] in their study of 144 patients aged >65 years who underwent urological surgery under spinal anesthesia, compared the median, paramedian, and ultrasound-assisted paramedian approaches. Their results demonstrated higher first-attempt success rates and fewer punctures in the ultrasound-assisted paramedian group. Although the paramedian approach used in their study offers advantages, such as a wider interlaminar space and avoidance of the supraspinous and interspinous ligaments, their findings are consistent with ours. Srinivasan et al.[13] and Schnabel et al.[14] also demonstrated that the use of ultrasound prior to spinal anesthesia significantly increases the incidence of successful first-attempt punctures.

In a study by Abdelhamid and Mansour,[15] the effect of ultrasound guidance on the ease of neuraxial anesthesia administration was examined in 90 ASA I-II patients with a mean age of 34.7 years using a midline approach. Their results, compared with our study, which included an older patient population, showed a first-attempt success rate of 80% in the ultrasound group and 37.8% in the surface landmark group. Additionally, the number of needle redirections was significantly lower in the ultrasound group. Unlike the study by Abdelhamid and Mansour,[15] our study additionally evaluated successful spinal anesthesia achieved with a single puncture and without needle redirection. In our study, both the first-attempt success rate and the proportion of procedures completed with a single puncture and no needle redirection were higher in the preprocedural ultrasound-guided marking group. An important consideration when interpreting our findings is the definition of the time-related variables. In this study, the recorded time reflects the spinal needle placement phase (from needle insertion to CSF flow and completion of the intrathecal injection) rather than the entire preprocedural workflow. Preprocedural ultrasound scanning and sterile preparation were performed before needle insertion and were not included in the recorded time. Therefore, the shorter duration observed in the ultrasound group represents improved technical performance during needle placement rather than a reduction in total clinical preparation time. Li et al.[9] performed spinal anesthesia in obese patients in the lateral position who were scheduled for cesarean delivery, and divided them into two BMI-based subgroups: 30-34.9 kg/m^2 and 35-43 kg/m^2. They found that, in the lower-BMI group, spinal needle-placement time was longer with ultrasound, whereas in the higher-BMI group it was shorter with ultrasound than with manual palpation. As BMI increases, palpation becomes more difficult, making the advantages of ultrasound more pronounced. Chin et al.[16] compared pre-procedural ultrasound with manual palpation for spinal anesthesia in 120 patients undergoing orthopedic surgery with difficult surface landmarks. They observed that although the time spent locating the puncture site was longer in the ultrasound group, the shorter duration of the spinal anesthesia procedure compensated for this difference, resulting in no overall disadvantage in spinal needle placement time compared with that in the manual palpation group. Abdelhamid and Mansour[15] evaluated the needle-placement phase from needle insertion to intrathecal drug administration and found this interval to be shorter in the ultrasound group. However, when the time required for preprocedural ultrasound scanning and interspinous-space marking was included, Abdelhamid and Mansour,[15] as well as Srinivasan et al.,[13] reported a longer total procedural duration in the ultrasound group. In contrast, Geng et al.,[17] in their study of 200 geriatric patients undergoing lower extremity surgery under neuraxial block, did not include evaluation of the interspinous space in their time measurements, and they found that procedure duration was significantly shorter in the ultrasound group.

In a study of patients with challenging surface anatomical landmarks, Chin et al.[16] evaluated 120 orthopedic surgery cases in which spinous processes were difficult to palpate, and reported that the first-attempt success rate for spinal anesthesia was higher in the ultrasound group. In our study, among patients whose spinous processes could not be palpated, the ultrasound group included a significantly higher number of such patients; nevertheless, the ultrasound group had significantly fewer punctures and needle redirections, and shorter times from needle insertion to CSF flow and for spinal needle placement.

Studies on this topic in obese patients have predominantly focused on pregnant populations. In this patient group, ultrasound-guided approaches reduce the number of punctures and shorten the procedure duration in technically difficult cases.[9,18,19] A study conducted in 2011 reported that pre-procedural ultrasound evaluation of the lumbar region in morbidly obese pregnant patients accurately predicted needle length, thereby avoiding

unnecessary punctures, reducing procedure time, and lowering complication rates.[20] Similarly, in our study, among patients with BMI ≥30 kg/m², the number of punctures, the number of needle redirections, the time from needle insertion to CSF flow, and the spinal needle-placement time were significantly lower in the ultrasound group.

Multiple puncture attempts during spinal anesthesia can cause patient discomfort and increase the risk of spinal hematoma, post-dural puncture headache, and nerve injury. Accurate identification of the interspinous space is a crucial step for successful spinal anesthesia. This minimizes the number of punctures, thereby reducing the risks of spinal hematoma, post-dural puncture headache, neural trauma, and patient discomfort.[21,22] The traditional manual palpation technique relies on anatomical landmarks, and this blind approach may lead to complications because of variability in patient anatomy. In their study, Lim et al.[23] reported a significantly lower incidence of paresthesia in the ultrasound group. Similarly, Dhanger et al.[18] observed that both the number of patients experiencing paresthesia and the number with blood in the spinal needle were significantly lower in the ultrasound group. In our study, no significant difference in the occurrence of paresthesia was observed between the groups. However, the number of patients in whom blood was detected in the spinal needle was higher in the manual palpation group. The total number of patients with paresthesia in our study was very small, which may explain the lack of statistical significance. Studies with larger patient populations might yield more meaningful results. In our study, the need for sedation during spinal anesthesia procedures was significantly lower in the ultrasound group. The higher sedation requirement in the manual palpation group may be related to the longer procedure duration and greater number of puncture attempts and needle redirections. Although paresthesia was numerically more frequent in the manual palpation group, the difference was not statistically significant.

Study Limitations

Our study has several limitations. First, the data were collected retrospectively, which precluded a randomized controlled design. A larger sample size could have provided more robust results, particularly in subgroup analyses of patients with high BMI or non-palpable spinous processes. Another limitation was the involvement of multiple practitioners in the administration of spinal anesthesia. However, there were no statistically significant differences in practitioner experience levels, and their impact on study outcomes was likely minimal. The retrospective design prevented the assessment of patient satisfaction, which could have provided further insights. During spinal anesthesia procedures, the doses of midazolam, fentanyl, or both administered for sedation in both groups were determined at the practitioner's discretion without using a standardized sedation scale. These doses were obtained from anesthesia records, which may have introduced variability. In our study, we used pre-procedural neuraxial ultrasound to identify the spinal anesthesia puncture site and compared this with the traditional manual palpation method. Although real-time ultrasound imaging during neuraxial anesthesia allows visualization of the entire needle trajectory during advancement, this approach is technically more challenging and often requires two operators. Patients who required conversion to general anesthesia were excluded, which may have led to a relative overestimation of procedural success. In addition, long-term outcomes were not available because of the retrospective design and the focus on peri-procedural variables.

CONCLUSION

Despite these limitations, our findings suggest that preprocedural ultrasound-guided marking is associated with improved first-attempt success and reduced technical difficulty during spinal anesthesia in elderly patients. This approach may be particularly useful in patients with difficult anatomical landmarks. However, prospective randomized studies with multivariable adjustment are needed to confirm whether this association reflects an independent effect.

Ethics

Ethics Committee Approval: This retrospective study was conducted at Dokuz Eylül University Faculty of Medicine Hospital after obtaining approval from the Non-Interventional Ethics Committee (approval number: 2018/17-36, date: 18.06.2019).

Informed Consent: Written informed consent for the anesthesia procedures was obtained from all patients as part of routine clinical care. No additional study-specific informed consent was obtained because of the retrospective design of the study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: H.K., H.A.E., Concept: H.K., H.A.E., S.K., Design: H.K., H.A.E., S.K., Data Collection or Processing: H.K., H.A.E., Analysis or Interpretation: H.K., H.A.E., Literature Search: H.K., Writing: H.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Biboulet P, Jourdan A, Van Haevre V, Morau D, Bernard N, Bringuier S, et al. Hemodynamic profile of target-controlled spinal anesthesia compared with 2 target-controlled general anesthesia techniques in elderly patients with cardiac comorbidities. Reg Anesth Pain Med. 2012;37:433-40.
2. Maxwell BG, Spitz W, Porter J. Association of increasing use of spinal anesthesia in hip fracture repair with treating an aging patient population. JAMA Surg. 2020;155:167-8.
3. Zeng W, Shi Y, Zheng Q, Du S. Ultrasound-assisted modified paramedian technique for spinal anesthesia in elderly. BMC Anesthesiol. 2022;22:242.
4. Papadakis M, Sapkas G, Papadopoulos EC, Katonis P. Pathophysiology and biomechanics of the aging spine. Open Orthop J. 2011;5:335-42.
5. Sahin T, Balaban O. Lumbar ultrasonography for obstetric neuraxial blocks: sonoanatomy and literature review. Turk J Anaesthesiol Reanim. 2018;46:257-67.
6. Shaikh F, Brzezinski J, Alexander S, et al. Ultrasound imaging for lumbar punctures and epidural catheterisations: systematic review and meta-analysis. BMJ. 2013;346:f1720.
7. Perlas A, Chaparro LE, Chin KJ. Lumbar neuraxial ultrasound for spinal and epidural anesthesia: a systematic review and meta-analysis. Reg Anesth Pain Med. 2016;41:251-60.
8. Khan MA, Gupta M, Sharma S, Kasaudhan S. A comparative study of ultrasound assisted versus landmark technique for combined spinal-epidural anaesthesia in patients undergoing lower limb orthopaedic surgery. Indian J Anaesth. 2022;66:272-7.
9. Li M, Ni X, Xu Z, et al. Ultrasound-assisted technology versus the conventional landmark location method in spinal anesthesia for cesarean delivery in obese parturients: a randomized controlled trial. Anesth Analg. 2019;129:155-61.
10. Li J, Krishna R, Zhang Y, Lam D, Vadivelu N. Ultrasound-guided neuraxial anesthesia. Curr Pain Headache Rep. 2020;24:59.
11. Park SK, Bae J, Yoo S, et al. Ultrasound-assisted versus landmark-guided spinal anesthesia in patients with abnormal spinal anatomy: a randomized controlled trial. Anesth Analg. 2020;130:787-95.
12. Michael J. Tessler, Ken Kardash, Richard M. Wahba, Simcha J. Kleiman, Stamatina T. Trihas, Michel Rossignol. The performance of spinal anesthesia is marginally more difficult in the elderly. Regional Anesthesia and Pain Medicine. 1999;24:126-30.
13. Srinivasan KK, Leo AM, Iohom G, Loughnane F, Lee PJ. Pre-procedure ultrasound-guided paramedian spinal anaesthesia at L5-S1: is this better than landmark-guided midline approach? A randomised controlled trial. Indian J Anaesth. 2018;62:53-60.
14. Schnabel A, Schuster F, Ermert T, Eberhart LH, Metterlein T, Kranke P. Ultrasound guidance for neuraxial analgesia and anesthesia in obstetrics: a quantitative systematic review. Ultraschall Med. 2012;33:E132-E7.
15. Abdelhamid SA, Mansour MA. Ultrasound-guided intrathecal anesthesia: does scanning help? Egyptian Journal of Anaesthesia. 2013;29:389-94.
16. Chin KJ, Perlas A, Chan V, Brown-Shreves D, Koshkin A, Vaishnav V. Ultrasound imaging facilitates spinal anesthesia in adults with difficult surface anatomic landmarks. Anesthesiology. 2011;115:94-101.
17. Geng J, Chen XL, Wang XD, Guo XY, Li M. [Ultrasound imaging increases first-attempt success rate of neuraxial block in elderly patients]. Zhonghua Yi Xue Za Zhi. 2016;96:3459-63.
18. Dhanger S, Vinayagam S, Vaidhyanathan B, Rajesh IJ, Tripathy DK. Comparison of landmark versus pre-procedural ultrasonography-assisted midline approach for identification of subarachnoid space in elective caesarean section: a randomised controlled trial. Indian J Anaesth. 2018;62:280-4.
19. Sahin T, Balaban O, Sahin L, Solak M, Toker K. A randomized controlled trial of preinsertion ultrasound guidance for spinal anaesthesia in pregnancy: outcomes among obese and lean parturients: ultrasound for spinal anesthesia in pregnancy. J Anesth. 2014;28:413-9.
20. Sahin T, Balaban O, Şahin B, Solak M, Toker K. Ultrasound guidance for subarachnoid block in parturient. Anestezi Dergisi. 2011;19:224-7.
21. Amin WA, Abou Seada MO, Bedair EMA, Elkersh MM, Karunakaran E. Comparative study between ultrasound determination and clinical assessment of the lumbar interspinous level for spinal anesthesia. Middle East J Anaesthesiol. 2014;22:407-12.
22. Reynolds F. Damage to the conus medullaris following spinal anaesthesia. Anaesthesia. 2001;56:238-47.
23. Lim YC, Choo CY, Tan KTJ. A randomised controlled trial of ultrasound-assisted spinal anaesthesia. Anaesth Intensive Care. 2014;42:191-8.

DOI: 10.4274/forbes.galenos.2026.57704
Forbes J Med 2026;7:84-89

Original Article

Predictors of Prolonged Hospitalization in Pediatric Respiratory Syncytial Virus Infections

Arzu ATEŞ[1], Mert KAMARA[2], Soner Sertan KARA[2]

[1]Nazilli State Hospital, Clinic of Clinical Biochemistry, Aydın, Türkiye
[2]Aydın Adnan Menderes University Faculty of Medicine, Department of Pediatric Infectious Diseases, Aydın, Türkiye

Cite as: Ateş A, Kamara M, Kara SS. Predictors of prolonged hospitalization in pediatric respiratory syncytial virus infections. Forbes J Med. 2026;7:84-89

ABSTRACT

Objective: Respiratory syncytial virus (RSV) is a leading cause of lower respiratory tract infections (LRTIs) in infants and young children. This study evaluated clinical and laboratory factors associated with prolonged hospitalization in pediatric RSV infection.

Methods: This retrospective study included children under 18 years of age who were hospitalized with LRTI and had confirmed RSV infection at Aydın Adnan Menderes University Hospital between 2023 and 2025. Demographic, clinical, and laboratory data were collected from hospital records. Patients were divided into two groups according to length of stay (LOS): ≤2 days and >2 days. Multivariate logistic regression analysis was performed using the backward stepwise method.

Results: A total of 185 RSV-infected children were analyzed. The median age was 2 months, and 51.9% were male. Patients with LOS >2 days had significantly higher rates of pediatric intensive care (PICU) admission, non-invasive ventilation, mechanical ventilation, and secondary bacterial infections. Laboratory findings in the prolonged-stay group showed higher lymphocyte counts and blood urea nitrogen (BUN) levels but lower mean platelet volume (MPV) levels. No mortality was observed. In the multivariate logistic regression analysis, PICU stay was associated with an approximately 12-fold increased risk of hospitalization lasting longer than 2 days.

Conclusion: Hospitalization longer than 2 days was associated with a more severe clinical course. In the multivariate logistic regression analysis, PICU stay was the main factor associated with hospitalization lasting longer than 2 days. Lymphocyte count, MPV, and BUN may provide supportive clinical information; however, their independent prognostic value remains uncertain.

Keywords: Respiratory syncytial virus, prolonged hospitalization, respiratory support, mean platelet volume

Received: 19.04.2026
Accepted: 21.07.2026
Epub: 21.07.2026
Publication Date: 13.08.2026

Corresponding Author:
Soner Sertan KARA
Aydın Adnan Menderes University Faculty of Medicine, Department of Pediatric Infectious Diseases, Aydın, Türkiye
drsoner@yahoo.com
ORCID ID: 0000-0002-8129-6063

INTRODUCTION

Respiratory syncytial virus (RSV) is among the most common etiologic agents of bronchiolitis and pneumonia, particularly in infants and older adults.[1] Lower respiratory tract infections (LRTIs) are a major cause of morbidity and mortality in children under 5 years of age, and RSV is one of the leading pathogens in this age group.[2,3] In temperate climates, RSV typically causes annual epidemics during autumn and winter, whereas in tropical regions, circulation often increases during the rainy season, and cases may occur year-round.[4,5] Globally, it was estimated that in 2019, RSV-associated acute LRTI occurred in 33.0 million children aged 0-60 months; these episodes resulted in 3.6 million hospitalizations, 26,300 in- hospital deaths, and a total of 101,400 deaths attributable to RSV.[6] In Türkiye, studies in children aged ≤2 years have shown high annual hospitalization incidences for LRTIs and bronchiolitis, with a substantial disease burden and higher RSV positivity clusters specifically in the 0-3 months age group.[7] Accordingly, RSV remains an important respiratory pathogen in both national and global clinical practice.

Severe RSV infection may present with hypoxemia and bronchiolitis-associated respiratory failure, particularly in young infants and children with significant comorbidities;

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

some cases require intensive care and invasive mechanical ventilation. A nationwide cohort study in Denmark reported that mechanical ventilation was required in 14.3 per 1000 RSV hospitalizations, primarily due to bronchiolitis-related respiratory failure in young infants and those with comorbidities, while extra-pulmonary complications were more prevalent among children without classic risk factors.[8]

In RSV infection, routine laboratory tests are not diagnostic; they are primarily used to support the clinical picture, monitor disease severity, and guide management when co-infection is suspected.

Reverse transcription polymerase chain reaction (RT-PCR) is the gold standard for diagnosis.[9] Regarding blood parameters, some studies have reported differences in C-reactive protein (CRP) and systemic inflammatory markers in RSV-positive cases; however, they are useful as adjuncts to clinical assessment rather than for confirming RSV infection.[10] In severe RSV bronchiolitis, hyponatremia associated with increased antidiuretic hormone (ADH) secretion/syndrome of inappropriate ADH and aspartate aminotransferase (AST)/alanine aminotransf (ALT) elevations may be observed.[11,12]

This study aimed to describe the clinico-laboratory characteristics of RSV-positive children and identify factors associated with clinical outcomes, specifically length of hospital stay and disease severity.

METHODS

This retrospective observational study identified 985 pediatric patients who were followed in the pediatric infectious diseases ward and/or the pediatric intensive care unit (PICU) for acute LRTI and tested for RSV between January 1, 2022, and December 31, 2025, via the Aydın Adnan Menderes University Hospital electronic database. Patients aged <18 years who were hospitalized for acute LRTI with confirmed RSV positivity in a respiratory specimen were included.

Exclusion criteria were: suspected nosocomial RSV infection (symptom onset ≥48-72 hours after admission), conditions markedly affecting laboratory analyses (e.g., active malignancy, chemotherapy, severe immunosuppression, advanced chronic kidney or liver disease), and missing baseline laboratory or primary outcome data.

All data were obtained from the Aydın Adnan Menderes University Hospital information system. Recorded variables included demographics (age, sex), clinical outcomes (length of hospital and PICU stay, need for non-invasive and invasive ventilation, secondary bacterial infections), and laboratory parameters [white blood cell count (WBC), neutrophils, lymphocytes, hemoglobin, platelets, mean platelet volume (MPV), Na, blood urea nitrogen (BUN), creatinine, AST, ALT, gamma-glutamyl transferase (GGT), albumin, and CRP]. The CRP/albumin ratio was calculated for the same time point.

RSV was detected using multiplex real-time RT-PCR on nasopharyngeal swabs or aspirates as part of routine clinical practice. Patients with two or more viral detections on the respiratory panel were excluded to avoid difficulties in interpretation. Secondary bacterial infection was defined as a clinically suspected bacterial infection recorded during the hospitalization period and supported by microbiological evidence and/or compatible radiological and laboratory findings. Bacterial detection by PCR was considered clinically significant only when accompanied by compatible clinical findings such as persistent or recurrent fever, increased inflammatory markers, radiological evidence of pneumonia, positive blood culture, or an antibiotic treatment decision by the attending physician. Because of the retrospective design, the exact timing of secondary bacterial infection in relation to admission could not be consistently determined in all cases. Therefore, secondary bacterial infection was not included in the multivariable model as a baseline predictor.

To identify factors associated with prolonged hospitalization, the study population was dichotomized according to length of stay (LOS) as ≤2 days and >2 days. Because the median LOS of the cohort was 2 days, this cutoff was selected to identify patients hospitalized for longer than the median. This median-based cutoff is consistent with previous pediatric RSV/acute lower respiratory infection research, in which longer LOS was defined as >2 days (the median).[13]

The study protocol was approved by the Aydın Adnan Menderes University Ethics Committee on April 14, 2026, with protocol number 2026/133.

Statistical Analysis

The Statistical Package for Social Sciences (SPSS) version 22 was used for statistical analyses (IBM SPSS Statistics for Windows, Version 22.0, NY, USA). Descriptive statistics were presented as number and percentage for normally distributed categorical data, while continuous variables were presented as median (25^{th}-75^{th} percentile) for non-normally distributed data. The normality of continuous variables was evaluated using the Kolmogorov-Smirnov test. Between-group comparisons were performed using the Mann-Whitney U test for continuous variables and the chi-square test or Fisher's exact test for categorical variables. Statistical significance was defined as $p<0.05$. Multivariate logistic regression analysis was performed using the backward stepwise method to evaluate factors associated with hospitalization lasting longer than 2

days. PICU stay, non-invasive ventilation, mechanical ventilation, lymphocyte count at diagnosis, BUN, and MPV were included in the initial model. Odds ratios and 95% confidence intervals (CIs) were calculated for the variables retained in the final model.

RESULTS

A total of 185 (18.8%) patients were diagnosed with RSV via multiplex real-time RT-PCR. Of these, 123(66.5%) had a LOS of ≤2 days and 62 (33.5%) had a stay of >2 days (Table 1). The median age was 2 months [interquartile range (IQR): 2-7] in the ≤2-day group and 3 months (IQR: 2-6) in the >2-day group, with no significant difference (p=0.915). Male sex accounted for 54.5% (67/123) of the ≤2-day group and 46.8% (29/62) of the >2-day group (p=0.323).

The median LOS was 1 day (IQR: 1-2) in the ≤2-day group and 3 days (IQR: 3-7) in the >2-day group (p<0.001). PICU admission was required for 14.1% (26/185) of all patients, differing significantly between groups: 4.1% (5/123) in the ≤2-day group versus 33.9% (21/62) in the >2-day group (p<0.001). Non-invasive ventilation was required in 7.6% (14/185) of patients and was more frequent in the >2-day group (19.4% vs. 1.6%, p<0.001). Invasive mechanical ventilation was required in 2.2% (4/185) of patients and occurred exclusively in the >2-day group (6.5% vs. 0%, p=0.012). Secondary bacterial infections were detected in 33.5% (62/185) of the cohort, with a significant difference between groups (p<0.001).

Complete blood count parameters (WBC, neutrophils, hemoglobin, and platelets) were comparable between groups, with the exception of lymphocytes and MPV. The median MPV was significantly higher in the ≤2-day group [10 (9-11) vs. 9 (9-10) fL, p=0.039]. Conversely, the median lymphocyte count was lower in the ≤2-day group [2349 (1314-3842) vs. 2891 (2000-4600), p=0.011]. Other laboratory parameters, including sodium, CRP, albumin, CRP/albumin ratio, AST, ALT, GGT, and creatinine, did not differ significantly between groups, except for BUN. Median

Table 1. Characteristics of the patients according to the duration of hospitalization

	Hospitalization ≤2 days n=123 (66.5%)	Hospitalization >2 days n=62 (33.5%)	p
Age, months, mean (25-75 percentile)	2 (2-7)	3 (2-6)	0.915
Male gender, n (%)	67 (54.5)	29 (46.8)	0.323
PICU stay, n (%)	5 (4.1)	21 (33.9)	**<0.001**
Non-invasive ventilation, n (%)	2 (1.6)	12 (19.4)	**<0.001**
Mechanical ventilation, n (%)	0 (0)	4 (6.5)	**0.012**
Secondary bacterial infections, n (%)	0 (0)	62 (33.5)	**<0.001**
Hyponatremia (serum Na <135 Meq/mL), n (%)	16 (13)	6 (9.7)	0.509
WBC, /mm^3, mean (25-75 percentile)	9536 (8353-11734)	10239 (8656-12000)	0.128
Neutrophils, /mm^3, mean (25-75 percentile)	5942 (4000-7245)	5967 (4200-6738)	0.578
Lymphocytes, /mm^3, mean (25-75 percentile)	2349 (1314-3842)	2891 (2000-4600)	**0.011**
Hemoglobin, gr/dL, mean (25-75 percentile)	11 (10-12)	11 (10-12)	0.556
Platelets, /mm^3, mean (25-75 percentile)	432000 (378000-481000)	395000 (365000-473000)	0.538
MPV, fL, mean (25-75 percentile)	10 (9-11)	9 (9-10)	**0.039**
Na, gr/dL, mean (25-75 percentile)	136 (136-137)	136 (136-138)	0.097
CRP, gr/dL, mean (25-75 percentile)	4 (1-8)	5 (1-12)	0.515
Albumin, gr/dL, mean (25-75 percentile)	39 (36-41)	38 (37-40)	0.849
AST, U/L, mean (25-75 percentile)	35 (27-46)	38 (28-50)	0.524
ALT, U/L, mean (25-75 percentile)	28 (22-34)	27.5 (22-36)	0.769
GGT, U/L, mean (25-75 percentile)	28 (21-39)	27.5 (20-41)	0.823
BUN, mg/dL, mean (25-75 percentile)	4 (4-8)	7 (4-9)	**0.023**
Creatinine, mg/dL, mean (25-75 percentile)	1 (1-1)	1 (1-1)	0.314
CRP/Albumin, mean (25-75 percentile)	0.10 (0.02-0.20)	0.12 (0.02-0.46)	0.551

PICU: Pediatric intensive care unit, Na: Sodium, WBC: White blood cell count, MPV: Mean platelet volume, CRP: C-reactive protein, AST: Aspartate aminotransferase, ALT: Alanine aminotransferase, GGT: Gamma-glutamyl transferase, BUN: Blood urea nitrogen, U/L: Units per liter

BUN was 4 mg/dL (4-8) in the ≤2-day group compared to 7 mg/dL (4-9) in the >2-day group (p=0.023).

In a logistic regression analysis for hospitalization longer than 2 days, performed using the backward stepwise method with PICU stay, non-invasive ventilation, mechanical ventilation, lymphocyte count at diagnosis, BUN, and MPV as candidate variables, PICU stay increased the risk by 12-fold in the final significant model (95% CI: 4.2-34.3) (Table 2).

DISCUSSION

In the present cohort, prolonged hospitalization due to RSV infection was associated with a more severe and clinically demanding disease course, suggesting a bidirectional relationship. Patients hospitalized for more than two days exhibited significantly higher rates of both non-invasive and invasive mechanical ventilation and a higher frequency of secondary bacterial infections. In addition, they had higher lymphocyte counts, lower MPV, and higher BUN.

The close relationship between longer LOS and greater respiratory support requirements has been demonstrated in previous studies. In the prospective National BRICK study, invasive mechanical ventilation rate was 61.1% among infants admitted to the PICU for RSV, indicating the high burden of respiratory support in this population.[14] In a Hungarian study of hospitalized children with severe acute respiratory infection, the median LOS was 5 days, and multivariable logistic regression showed that receipt of mechanical ventilation or oxygen support increased the likelihood of hospitalization lasting >4 days , pneumonia, and RSV etiology by 3.23-, 3.65-, and 3.25-fold, respectively.[15] Similarly, in another study of RSV-associated LRTI cases, "severe RSV" was defined as any of the following: hospitalization for >7 days, ICU admission, non-invasive or invasive ventilation, or in-hospital mortality; 40.16% of cases were reported to meet this definition.[16] In our study, hospitalization lasting longer than 2 days was used as a median-based outcome measure to evaluate factors associated with a more severe clinical course among children with RSV infection.

The association between secondary bacterial infection and prolonged hospitalization is clinically plausible. Because the timing of secondary bacterial infection could not be consistently determined, this association should be interpreted as a disease-course relationship rather than as evidence of baseline prediction. Recent studies have similarly shown that bacterial infections accompanying RSV can increase both disease severity and length of hospitalization. In a multicenter prospective surveillance study in Portugal, potentially pathogenic bacteria were detected in 20.4% of hospitalized cases of RSV; in this group, PICU admission was more frequent, hospital stays were longer, and the risk of severe disease was reported to be increased 13.2-fold.[17] Across different series, bacterial coinfection was identified in 32.4% of children with RSV pneumonia and 47.3% of hospitalized RSV-positive children. Pneumococcus was the most prevalent pathogen, and children with bacterial coinfection had longer hospital stays and a greater need for PICU care.[18] LOS remained one of the independent predictors of bacterial coinfection.[19] Diaz-Diaz et al.[20] defined a prolonged stay as lasting longer than 3 days. They described nasopharyngeal coexistence of *Streptococcus pneumoniae* and/or *Haemophilus influenzae* as being associated with a more severe clinical course, particularly increasing the likelihood of requiring oxygen and of prolonged hospitalization.

In this study, prolonged hospitalization due to RSV infection was associated with higher lymphocyte counts. In contrast, the study by Matera et al.[21] involving 1,297 infants hospitalized for bronchiolitis reported that RSV positivity was more frequent in the lowest lymphocyte group, which was associated with higher clinical severity scores, greater oxygen requirements, longer hospital stays, and more frequent PICU unit admissions. In another study that evaluated 325 children with RSV-associated bronchiolitis, each unit increase in lymphocyte percentage was associated with a lower likelihood of severe disease, similar to our study.[22] The lymphocyte difference observed in our study is not merely an isolated laboratory variable but may represent an immunological response associated with a more severe clinical course and longer hospitalization.

We observed lower MPV levels in patients with prolonged hospitalization, which may reflect platelet consumption or an altered platelet response during a more severe and

Table 2. Multivariate logistic regression analysis of risk factors for prolonged hospitalization

Independent variable	Odds ratio	95% CI	p
PICU admission	12.047	4.236-34.263	<0.001
MPV	0.746	0.527-1.056	0.098

Backward stepwise logistic regression was performed. Nagelkerke R^2 =0.219; Cox and snell R^2 =0.158; Hosmer-Lemeshow $\chi^2(4)=6.712$
PICU: Pediatric intensive care unit, MPV: Mean platelet volume, CI: Confidence interval

sustained inflammatory process. However, MPV does not show a consistent trend across RTIs. While Doğan et al.[23] reported increased MPV associated with bronchiolitis severity in 179 cases, Feketea et al.[24] found lower MPV in non-influenza viral respiratory infections compared to influenza and healthy controls. To date, no study has specifically evaluated MPV in RSV. Although our findings align with some data from viral RTIs, further studies are needed before MPV can be considered an RSV-specific marker or a reliable indicator of disease severity.

We found higher BUN levels in the prolonged-stay group, which likely reflect reduced oral intake, dehydration, and increased catabolic stress rather than intrinsic renal dysfunction, since creatinine levels did not differ between groups. This is indirectly supported by the bronchiolitis literature, which identifies dehydration and poor feeding—conditions often accompanied by clinically significant fluid-electrolyte disturbances—as predictors of escalated care. However, specific evidence supporting BUN as an independent severity marker for RSV remains limited, and recent predictive studies have failed to identify urea as an independent risk factor. Consequently, elevated BUN should be regarded as a supportive clinical finding rather than a definitive prognostic marker.[25-27]

Apart from the variables identified as significant, most clinical and laboratory parameters did not differ between groups defined by length of hospital stay. Patients with shorter and prolonged hospitalizations were similar with respect to age, sex, hyponatremia, total WBC and neutrophil counts, hemoglobin, platelet counts, sodium, CRP, albumin, AST, ALT, GGT, creatinine, and the CRP/albumin ratio. These findings suggest that LOS-related differences were limited to selected clinical severity indicators and specific laboratory parameters, while the majority of baseline characteristics remained comparable between groups.

Although this study provides clinically relevant data on factors associated with prolonged RSV hospitalization, notably respiratory support requirements, secondary bacterial infections, and specific laboratory abnormalities, it has several limitations. First, the single-center retrospective design is inherently subject to selection bias and has limited generalizability. Important clinical confounders such as prematurity, congenital heart disease, chronic lung disease, immunodeficiency, and other underlying medical conditions could not be consistently retrieved from the retrospective database. The absence of these variables may have resulted in residual confounding and may limit the causal interpretation of the observed associations. Clinical decisions, such as PICU transfers and initiation of ventilation, were based on physicians' judgment rather than a standardized protocol, potentially introducing practice-based variability. Second, dichotomizing LOS (≤2 vs. >2 days) simplified the analysis but may have reduced data granularity and obscured dose-response relationships. Third, the lack of consistent data on ventilation duration and dynamic severity markers limited the ability to more comprehensively characterize the disease burden. Finally, potentially repeated admissions of the same patients may have influenced the frequencies of clinical outcomes.

CONCLUSION

In this pediatric RSV cohort, hospitalization longer than 2 days was associated with a more severe clinical course, including PICU admission and greater need for respiratory support. In the multivariate logistic regression analysis, PICU stay emerged as the main factor associated with hospitalization lasting longer than 2 days. Secondary bacterial infections were more frequent in the prolonged-stay group; however, because their timing could not be consistently determined, this finding should be interpreted cautiously as part of the clinical disease course rather than as a baseline predictor. Selected laboratory parameters, including lymphocyte count, MPV, and BUN, may provide supportive clinical information; however, their independent prognostic value remains uncertain. Prospective studies including major clinical confounders and validated LOS thresholds are needed to confirm these findings.

Ethics

Ethics Committee Approval: The study protocol was approved by the Aydın Adnan Menderes University Ethics Committee on April 14, 2026, with protocol number 2026/133.

Informed Consent: This retrospective study was conducted using data obtained from existing hospital records.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.A., S.S.K., Concept: A.A., S.S.K., Design: A.A., S.S.K., Data Collection or Processing: A.A., M.K., S.S.K., Analysis or Interpretation: A.A., M.K., S.S.K., Literature Search: A.A., M.K., S.S.K., Writing: A.A., M.K., S.S.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Qiu X, Xu S, Lu Y, et al. Development of mRNA vaccines against respiratory syncytial virus (RSV). Cytokine Growth Factor Rev. 2022;68:37-53.

2. GBD 2016 Lower Respiratory Infections Collaborators. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet Infect Dis. 2018;18:1191-210.
3. 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;392:1736-88. Erratum in: Lancet. 2019;393:e44. Erratum in: Lancet. 2018;392:2170.
4. Borchers AT, Chang C, Gershwin ME, Gershwin LJ. Respiratory syncytial virus--a comprehensive review. Clin Rev Allergy Immunol. 2013;45:331-79.
5. Obando-Pacheco P, Justicia-Grande AJ, Rivero-Calle I, et al. Respiratory syncytial virus seasonality: a global overview. J Infect Dis. 2018;217:1356-64.
6. Li Y, Wang X, Blau DM, et al; Respiratory Virus Global Epidemiology Network; Nair H; RESCEU investigators. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. Lancet. 2022;399:2047-64.
7. Hacımustafaoğlu M, Celebi S, Bozdemir SE, et al. RSV frequency in children below 2 years hospitalized for lower respiratory tract infections. Turk J Pediatr. 2013;55:130-9.
8. Nygaard U, Hartling UB, Nielsen J, et al. Hospital admissions and need for mechanical ventilation in children with respiratory syncytial virus before and during the COVID-19 pandemic: a Danish nationwide cohort study. Lancet Child Adolesc Health. 2023;7:171-9.
9. Onwuchekwa C, Atwell J, Moreo LM, et al. Pediatric respiratory syncytial virus diagnostic testing performance: a systematic review and meta-analysis. J Infect Dis. 2023;228:1516-27.
10. Okuyan O, Elgormus Y, Dumur S, Sayili U, Uzun H. New generation of systemic inflammatory markers for respiratory syncytial virus infection in children. Viruses. 2023;15:1245.
11. Eisenhut M. Extrapulmonary manifestations of severe respiratory syncytial virus infection--a systematic review. Crit Care. 2006;10:R107.
12. Eisenhut M, Thorburn K, Ahmed T. Transaminase levels in ventilated children with respiratory syncytial virus bronchiolitis. Intensive Care Med. 2004;30:931-4.
13. Ozturk A, Chan M, Khan JR, et al. Clinical characteristics and in-hospital outcomes associated with respiratory syncytial virus vs other viral acute lower respiratory infections in hospitalized children younger than 2 years. J Infect Dis. 2025;231:1309-17.
14. Phijffer EWEM, Wildenbeest JG, Brouwer CNM, et al. The full burden of RSV-related pediatric intensive care unit admissions during infancy: a prospective national study (BRICK Study). Pediatr Infect Dis J. 2025;44:517-21.
15. Orosz N, Gömöri G, Battamir U, Nagy AC. Hospital-based cross-sectional study on the clinical characteristics of children with severe acute respiratory infections in Hungary. BMC Infect Dis. 2024;24:1268.
16. Bunjoungmanee P, Sompoch S, Tangsathapornpong A, Kulalert P. Factors associated with severe respiratory syncytial virus infection among hospitalized children in Thammasat University Hospital. F1000Res. 2024;13:231.
17. Torres AR, Gaio V, Melo A, et al; VigiRSV group. RSV-bacterial co-infection is associated with increased illness severity in hospitalized children - results from a prospective sentinel surveillance study. J Med Virol. 2025;97:e70209.
18. Lin HC, Liu YC, Hsing TY, et al. RSV pneumonia with or without bacterial co-infection among healthy children. J Formos Med Assoc. 2022;121:687-93.
19. Qu X, Ye X, Yu J, et al. Epidemiological and clinical characteristics of bacterial co-detection in respiratory syncytial virus-positive children in Wenzhou, China, 2021 to 2023. BMC Infect Dis. 2025;25:697.
20. Diaz-Diaz A, Bunsow E, Garcia-Maurino C, et al. Nasopharyngeal codetection of haemophilus influenzae and streptococcus pneumoniae shapes respiratory syncytial virus disease outcomes in children. J Infect Dis. 2022;225:912-23.
21. Matera L, Nenna R, Frassanito A, et al. Low lymphocyte count: a clinical severity marker in infants with bronchiolitis. Pediatr Pulmonol. 2022;57:1770-5.
22. Yan J, Zhao L, Zhang T, et al. Development and validation of a nomogram for predicting severe respiratory syncytial virus-associated bronchiolitis. BMC Infect Dis. 2023;23:249.
23. Doğan M, Köse M, Öztürk MA, Hangül M, Aslaner H. The relationship between immature platelet fraction and severity of acute bronchiolitis. Turk J Pediatr. 2021;63:1056-63.
24. Feketea G, Vlacha V, Pop RM, et al. Relationship between vitamin D level and platelet parameters in children with viral respiratory infections. Front Pediatr. 2022;10:824959.
25. Freire G, Kuppermann N, Zemek R, et al; Pediatric emergency research networks (PERN). Predicting escalated care in infants with bronchiolitis. Pediatrics. 2018;142:e20174253. Erratum in: Pediatrics. 2019;143:e20183404.
26. Poddar U, Singhi S, Ganguli NK, Sialy R. Water electrolyte homeostasis in acute bronchiolitis. Indian Pediatr. 1995;32:59-65.
27. Liu HF, Zhang XZ, Liu CY, et al. A novel combined nomogram for predicting severe acute lower respiratory tract infection in children hospitalized for RSV infection during the post-COVID-19 period. Front Immunol. 2024;15:1437834.

DOI: 10.4274/forbes.galenos.2026.92609
Forbes J Med 2026;7:90-7

Original Article

Early Predictors of Non-Invasive Ventilation Failure in COPD Patients with Acute Hypercapnic Respiratory Failure: A Multicenter Retrospective Study

Damla SERÇE UNAT[1], Nilay ERTEN TURAN[1], Ömer Selim UNAT[2], Mutlu Onur GÜÇSAV[1], Emine Sena DİKMENTEPE[3], Hakan KAHRAMAN[2], Aydan MERTOĞLU[2]

[1]İzmir Bakırçay University Faculty of Medicine, Department of Pulmonology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital, Clinic of Translational Pulmonology, İzmir, Türkiye
[3]Siirt Training and Research Hospital, Clinic of Pulmonology, Siirt, Türkiye

Cite as: Serçe Unat D, Erten Turan N, Unat ÖS, et al. Early predictors of non-invasive ventilation failure in COPD patients with acute hypercapnic respiratory failure: a multicenter retrospective study. Forbes J Med. 2026;7:90-7

ABSTRACT

Objective: Early identification of non-invasive ventilation (NIV) failure in patients with acute hypercapnic respiratory failure due to chronic obstructive pulmonary disease (COPD) exacerbation remains a critical clinical challenge, particularly in emergency departments. This study aimed to evaluate whether early arterial blood gas changes after NIV initiation predict intensive care unit (ICU) admission and in-hospital intubation and to develop a simple predictive model based on early physiological response.

Methods: This multicenter retrospective study included 319 patients with COPD exacerbation complicated by acute hypercapnic respiratory failure who were initiated on NIV in the emergency departments of three hospitals. Arterial blood gas parameters were recorded before NIV initiation and 1-3 hours thereafter. Early response was defined as the change in $PaCO_2$ ($\Delta PaCO_2$). Multivariable logistic regression models were constructed for ICU admission and intubation. Model discrimination was assessed using receiver operating characteristic curve analysis.

Results: During hospitalization, 138 patients (43.3%) required ICU admission, and 70 patients (21.9%) required intubation. Lower post-NIV pH, insufficient reduction in $\Delta PaCO_2$, and prior ICU admission independently predicted the need for ICU care and intubation. The model demonstrated good discriminative performance for ICU admission [area under the curve (AUC) =0.787; 95% confidence interval (CI): 0.737-0.837] and for intubation (AUC =0.810; 95% CI: 0.746-0.873).

Conclusion: Early physiological response to NIV, particularly post-treatment pH and the magnitude of PaCO2 reduction, provides clinically meaningful prognostic information and may assist early decision-making in emergency settings. External validation is warranted.

Keywords: Pulmonary disease, chronic obstructive lung disease, respiratory insufficiency, noninvasive ventilation, intubation, intensive care units

Received: 16.05.2026
Accepted: 03.08.2026
Epub: 05.08.2026
Publication Date: 20.08.2026

Corresponding Author:
Damla SERÇE UNAT
İzmir Bakırçay University Faculty of Medicine, Department of Pulmonology, İzmir, Türkiye
sercedamla@gmail.com
ORCID ID: 0000-0003-4743-5469

INTRODUCTION

Chronic obstructive pulmonary disease (COPD) is a progressive disorder characterized by persistent airflow limitation and represents a major public health problem associated with substantial morbidity and mortality. According to the World Health Organization, COPD is the third leading cause of death worldwide, accounting for more than 3 million deaths annually.[1] Global Burden of Disease analyses indicate that the disease burden continues to rise, particularly in low- and middle-income countries.[2] In Türkiye, the BOLD study reported a spirometry-confirmed COPD prevalence of approximately 19% among adults aged over 40 years.[3] Acute exacerbations of COPD are associated with increased hospitalizations and mortality, and each exacerbation contributes to disease progression and clinical deterioration.[4]

Severe exacerbations may lead to acute hypercapnic respiratory failure as a consequence of alveolar

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

hypoventilation and respiratory muscle fatigue, significantly increasing the risk of mortality.[5,6] In this setting, non-invasive mechanical ventilation (NIV) is recommended as the first-line ventilatory support modality due to its ability to improve gas exchange, reduce the work of breathing, and decrease the need for endotracheal intubation.[7-9] However, NIV failure has been reported in approximately 15-30% of cases, and delayed recognition of treatment failure may further increase mortality.[10,11] This creates a critical clinical decision-making challenge, particularly in emergency departments, where timely intubation decisions are essential.

Among patients initiated on NIV in the emergency department, early clinical and arterial blood gas responses play a decisive role in determining whether to continue NIV or to escalate to invasive mechanical ventilation. Previous studies have demonstrated that persistent acidosis and insufficient reduction in $PaCO_2$ during the first hours of NIV are associated with treatment failure.[12,13] Nevertheless, most available data are derived from intensive care unit (ICU) populations, and evidence focusing specifically on the early response to NIV initiated in the emergency department remains limited.

The present study aimed to evaluate the association between early clinical and arterial blood gas parameters and clinical outcomes in patients with acute hypercapnic respiratory failure due to COPD exacerbation who were started on NIV in the emergency department and to identify variables that may predict the need for intubation.

METHODS

Study Design and Setting

This multicenter retrospective observational study was conducted across three hospitals in Türkiye: a secondary-level state hospital, this multicenter retrospective observational study was conducted across three hospitals in Türkiye: Giresun Dr. Ali Menekşe Chest Diseases Hospital, a secondary care hospital; İzmir Bakırçay University, a tertiary multidisciplinary training and research hospital; and University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital, a tertiary hospital specialized in chest diseases. Patients presenting to the emergency departments with acute hypercapnic respiratory failure due to COPD exacerbation and who were initiated on NIV were included. Differences in data collection periods across centers were due to institutional transitions involving members of the research team. The study was reported in accordance with the STROBE statement for observational studies.

Ethics Statement

This study was approved by the Clinical Research Ethics Committee of İzmir Bakırçay University Faculty of Medicine (approval no: 2702, date: 04.03.2026). The study was conducted in accordance with the principles of the Declaration of Helsinki. Institutional permissions were obtained from all three participating centers, and administrative approval for data use was granted by the respective hospital authorities. Due to the retrospective design and use of anonymized data, the requirement for informed consent was waived by the Ethics Committee.

Patient Selection

All patients presenting to the emergency department with a diagnosis of COPD exacerbation complicated by acute hypercapnic respiratory failure and who were initiated on NIV were screened for eligibility. COPD exacerbation was defined clinically as worsening dyspnea and/or changes in sputum volume or purulence requiring acute medical treatment. Acute hypercapnic respiratory failure was diagnosed based on baseline arterial blood gas analysis demonstrating hypercapnia accompanied by respiratory acidosis.

Patients were included if NIV was initiated in the emergency department, a baseline arterial blood gas measurement was obtained prior to NIV initiation, and a follow-up arterial blood gas measurement was available within 1-3 hours after NIV initiation. Patients were excluded if only venous blood gas analysis was available, if the follow-up arterial blood gas measurement was missing, if cardiac arrest occurred within the first 2 hours after NIV initiation, or if endotracheal intubation was required within the first 2 hours after NIV initiation.

When multiple arterial blood gas measurements were available within the 1-3-hour window, the measurement closest to 2 hours after NIV initiation was selected for analysis to ensure consistency in the assessment of early response.

NIV Application and Monitoring

NIV was delivered via a face mask using bilevel positive airway pressure (BPAP). Initial ventilator settings were determined by the treating physician based on clinical status and patient tolerance. Patients were closely monitored for respiratory effort, oxygenation, mental status, and hemodynamic stability. Intubation decisions were made according to standard clinical criteria, including deterioration in level of consciousness, hemodynamic instability, refractory hypercapnia/hypoxemia, or NIV intolerance.

Data Collection and Variables

Data collection focused on demographic and clinical characteristics, including comorbidities, history of ICU admission and intubation, use of home oxygen and BPAP, and hospitalization within the previous year. Laboratory variables included hemoglobin, white blood cell count, and C-reactive protein (CRP). Arterial blood gas measurements (pH, $PaCO_2$, PaO_2, and HCO_3-) were obtained at baseline (before NIV initiation) and repeated 1-3 hours after NIV initiation. Early physiological response to NIV was quantified as the change in $PaCO_2$ over this interval and calculated as $\Delta PaCO_2$ = $PaCO_2$ (post-NIV)-$PaCO_2$ (baseline).

Two primary outcomes were prespecified: (i) requirement for ICU admission at any time during the index hospitalization; and (ii) requirement for endotracheal intubation during the index hospitalization. For the purposes of this study, NIV failure was defined as the need for endotracheal intubation and transition to invasive mechanical ventilation during the index hospitalization.

Statistical Analysis

Continuous variables were assessed for normality and expressed as mean ± standard deviation or median (interquartile range). Group comparisons were performed using Student's t-test, Mann-Whitney U test, and chi-square test, as appropriate. Variables that were significant in univariate analysis (p<0.10) or that were clinically relevant were entered into multivariable logistic regression models. Separate models were constructed for ICU admission and intubation.

Model Development and Probability Estimation

Multivariable logistic regression coefficients were used to estimate each patient's predicted probability of the outcome. The optimal probability threshold was identified from the receiver operating characteristic (ROC) curve using the Youden index, and a spreadsheet-based calculator was prepared to facilitate bedside application.

Model Performance

Discrimination was evaluated using ROC curve analysis and the area under the curve (AUC) with a 95% confidence interval (CI). At the optimal probability threshold identified by ROC analysis, the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) of each model were calculated. A p value <0.05 was considered statistically significant.

RESULTS

Study Population

A total of 319 patients were included from three centers: Giresun Dr. Ali Menekşe Chest Diseases Hospital (n=110), a secondary care hospital; İzmir Bakırçay University (n=109), a tertiary multidisciplinary training and research hospital; and University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital (n=100), a tertiary hospital specialized in chest diseases. During hospitalization, 138 patients (43.3%) required ICU admission, and 70 patients (21.9%) required endotracheal intubation.

Findings According to ICU Requirement

Patients requiring ICU admission were significantly older than those managed without ICU care (71.9±12.0 vs. 69.1±10.8 years; p=0.032). Baseline arterial blood gas analysis demonstrated lower pH levels in the ICU group (7.29±0.07 vs. 7.31±0.06; p=0.001), and post-NIV pH remained significantly lower in these patients (7.32±0.10 vs. 7.36±0.05; p<0.001).

The median baseline $PaCO_2$ level was higher in patients requiring ICU admission [69.5 mmHg (59.4-82.0) vs. 65.0 mmHg (55.1-71.2); p=0.001]. Similarly, post-NIV $PaCO_2$ values were significantly higher in the ICU group [65.5 mmHg (56.2-78.9) vs. 57.9 mmHg (49.4-64.4); p<0.001]. The early physiological response reflected by $\Delta PaCO_2$ was significantly smaller among patients who required ICU care (-3.51±16.88 mmHg vs. -8.30±10.39 mmHg; p=0.002).

CRP levels were markedly higher in the ICU group [47.5 mg/L (19.4-97.6) vs. 26.5 mg/L (8.1-67.9); p=0.003]. The length of hospital stay was also significantly longer in these patients [12.5 days (8-20) vs. 8 days (5-12); p<0.001]. Prior ICU admission (63.8% vs. 23.3%), previous intubation (26.1% vs. 3.3%), and hospitalization within the previous year (97.1% vs. 82.2%) were strongly associated with ICU requirement (all p<0.001) (Table 1).

In multivariable logistic regression analysis, lower post-NIV pH, insufficient reduction in $\Delta PaCO_2$, and prior ICU admission remained independent predictors of ICU requirement. The model demonstrated good discriminative ability (AUC =0.787; 95% CI: 0.737-0.837; p<0.001). The ROC curve is presented in Figure 1 .At the optimal predicted probability cut-off value of 0.430, the model predicted ICU requirement with a sensitivity of 67.4%, specificity of 76.7%, PPV of 68.9%, and NPV of 75.4%.

Findings According to Intubation Requirement

Patients who required intubation had significantly lower hemoglobin levels compared with those who did not [12.5 g/dL (11.2-14.5) vs. 13.3 g/dL (11.8-15.1); p=0.016]. CRP levels were significantly higher in the intubation group [55.0 mg/L (23.4-86.7) vs. 27.4 mg/L (8.1-68.1); p<0.001].

Baseline pH was significantly lower in patients who required intubation [7.28 (7.24-7.33) vs. 7.31 (7.27-7.35); p<0.001]; the

pH difference was more pronounced after NIV initiation [7.28 (7.22-7.34) vs. 7.37 (7.32-7.40); p<0.001]. Baseline $PaCO_2$ levels were higher in the intubation group [69.8 mmHg (60.1-90.5) vs. 65.7 mmHg (56.4-72.8); p=0.022], and post-NIV $PaCO_2$ remained markedly elevated [72.8 mmHg (60.1-95.4) vs. 58.9 mmHg (50.2-65.7); p<0.001]. The reduction in $PaCO_2$ ($\Delta PaCO_2$) was significantly smaller among patients who required intubation [-0.8 mmHg (-11.2 to 13.4) vs. -7.4 mmHg (-14.1 to -3.5); p<0.001].

Length of hospital stay was significantly longer in intubated patients [15.5 days (9-23) vs. 9 days (5-13); p<0.001]. Prior ICU admission (74.3% vs. 31.7%), previous intubation (40.0% vs. 6.0%), home oxygen use (80.0% vs. 66.7%), and hospitalization within the previous year (98.6% vs. 85.9%) were significantly associated with the requirement for intubation (p<0.05) (Table 2).

In multivariable analysis, lower post-NIV pH, insufficient reduction in $\Delta PaCO_2$, and prior ICU admission remained independent predictors of intubation. The model demonstrated good discriminative performance (AUC =0.810; 95% CI: 0.746-0.873; p<0.001), and the ROC curve is shown in Figure 2. At the optimal predicted probability cut-off value of 0.303, the model predicted the need for intubation with a sensitivity of 64.3%, specificity of 86.3%, PPV of 57.0%, and NPV of 89.6%.

Table 1. Comparison of clinical, laboratory and arterial blood gas variables according to ICU requirement

Parameters	No ICU (n=181)	ICU (n=138)	p value
Age, years	69.1±10.8	71.9±12.0	**0.032**
pH (baseline)	7.31±0.06	7.29±0.07	**0.001**
pH (post-NIV)	7.36±0.05	7.32±0.10	**<0.001**
HCO_3- (baseline), mmol/L	30.47±4.96	30.85±5.93	0.535
HCO_3- (post-NIV), mmol/L	30.86±4.76	32.74±10.85	**0.039**
PaO_2 (baseline), mmHg	63.5 (52.4-74.8)	63.5 (54.1-72.6)	0.576
$PaCO_2$ (baseline), mmHg	65.0 (55.1-71.2)	69.5 (59.4-82.0)	**0.001**
PaO_2 (post-NIV), mmHg	66.8 (55.3-80.6)	67.8 (54.9-84.1)	0.531
$PaCO_2$ (post-NIV), mmHg	57.9 (49.4-64.4)	65.5 (56.2-78.9)	**<0.001**
$\Delta PaCO_2$ (post-baseline), mmHg	-8.30±10.39	-3.51±16.88	**0.002**
Hemoglobin, g/dL	13.4 (11.9-14.9)	12.8 (11.1-14.4)	0.100
Leukocyte count, /mm^3	11.150 (7.800-13 760)	11.205 (7.200-14 800)	0.293
CRP, mg/L	26.5 (8.1-67.9)	47.5 (19.4-97.6)	**0.003**
Length of hospital stay, days	8 (5-12)	12.5 (8-20)	**<0.001**
Female sex, n (%)	56 (31.1)	41 (29.7)	0.788
Smoking status, n (%)			
Non-smoker	29 (16.1)	21 (15.2)	0.113
Ex-smoker	75 (41.7)	73 (52.9)	
Current smoker	76 (42.2)	44 (31.9)	
Hypertension, n (%)	103 (57.2)	91 (65.9)	0.114
Diabetes mellitus, n (%)	40 (22.2)	37 (26.8)	0.320
History of cerebrovascular disease, n (%)	5 (2.8)	7 (5.1)	0.287
Cardiovascular disease, n (%)	63 (35.0)	46 (33.3)	0.756
Previous ICU admission, n (%)	42 (23.3)	88 (63.8)	**<0.001**
History of intubation, n (%)	6 (3.3)	36 (26.1)	**<0.001**
Home oxygen therapy, n (%)	118 (65.6)	103 (74.6)	0.081
Home BPAP use, n (%)	72 (40.0)	45 (32.6)	0.176
Hospitalization within previous year, n (%)	148 (82.2)	134 (97.1)	**<0.001**

Data are expressed as mean ± standard deviation or median (25^{th}-75^{th} interquartile range) for continuous variables, and as number (percentage) for categorical variables. p values were calculated using appropriate parametric or non-parametric tests based on data distribution. Statistically significant p values are shown in bold. BPAP: Bilevel positive airway pressure, CRP: C-reactive protein, HCO_3-: Bicarbonate, ICU: Intensive care unit, NIV: Non-invasive ventilation, $PaCO_2$: Arterial partial pressure of carbon dioxide, PaO_2: Arterial partial pressure of oxygen, $\Delta PaCO_2$: Change in arterial carbon dioxide level calculated as post-NIV, $PaCO_2$ minus baseline $PaCO_2$

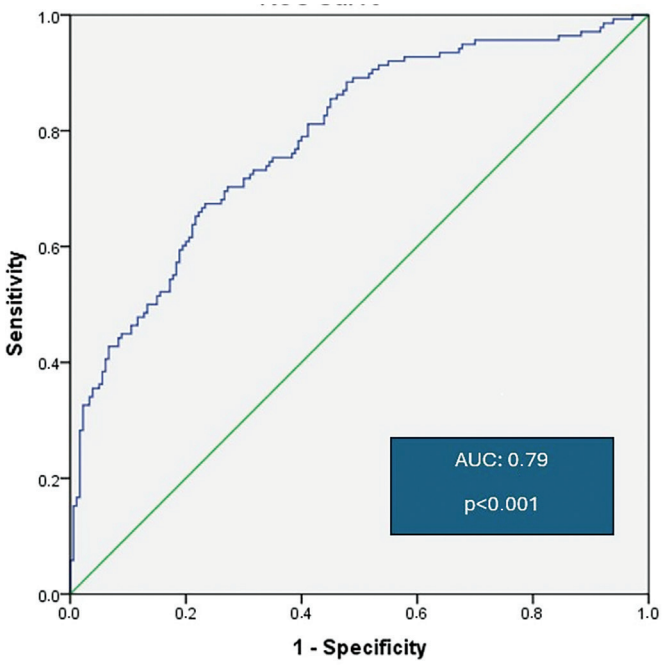


Figure 1. Receiver operating characteristic curve of the regression-based model predicting intensive care unit admission in patients with acute hypercapnic respiratory failure due to chronic obstructive pulmonary disease exacerbation treated with non-invasive ventilation in the emergency department. The model demonstrated good discriminative performance with an AUC of 0.787 (95% confidence interval: 0.737-0.837; p<0.001)

AUC: Area under the curve

DISCUSSION

In this multicenter cohort of patients with acute hypercapnic respiratory failure due to COPD exacerbation who were treated with NIV in the emergency department, early post-treatment physiological response emerged as the strongest determinant of clinical trajectory. Lower post-NIV pH, insufficient reduction in $PaCO_2$ within the first 1-3 hours, and a history of prior ICU admission independently predicted both ICU requirement and in-hospital intubation. The discriminative performance of the derived models was clinically meaningful, with AUC values of 0.787 for ICU admission and 0.810 for intubation. These findings suggest that early dynamic changes following NIV initiation provide prognostic information beyond baseline severity markers.

Several prognostic tools have been developed to predict mortality in acute exacerbations of COPD. The NIVO score estimates in-hospital mortality in patients requiring assisted ventilation, using admission variables such as age and acidosis.[14] Similarly, the DECAF score predicts hospital mortality based on clinical and laboratory parameters at presentation.[15] While these tools offer valuable risk stratification for mortality, they are primarily static and not specifically designed to predict NIV failure or early need for escalation. Our model differs by incorporating early treatment response variables, particularly post-NIV pH and $\Delta PaCO_2$, thus reflecting the physiological evolution under ventilatory support rather than admission status alone.

Beyond mortality-focused scores, efforts have been made to predict NIV failure directly. Duan et al.[16] developed and validated a simple early risk score for NIV failure in COPD patients, and van Gemert et al.[17] identified factors associated with intubation following NIV failure in the emergency department. Although these studies highlight the importance of early identification of high-risk patients, most models rely heavily on baseline characteristics. In contrast, our findings demonstrate that the magnitude of $PaCO_2$ reduction during the initial hours of therapy independently predicts deterioration, reinforcing the concept that response to treatment is as important as initial severity.

The efficacy of NIV in acute hypercapnic COPD has been well established since the landmark randomized trial by Brochard et al.,[18] which demonstrated reduced intubation rates and mortality. The benefits of early NIV initiation were further confirmed by Plant et al.,[19] even in non-ICU ward settings. More recent real-world data from large national cohorts have shown that, despite widespread NIV adoption, outcomes remain heterogeneous, and failure continues to occur in a significant proportion of patients.[20] Our study builds upon this foundation by addressing a clinically relevant but underexplored question: how to identify patients early who are unlikely to benefit from continued NIV. These observations are consistent with current guideline recommendations: the ERS/ATS clinical practice guidelines endorse NIV as the first-line ventilatory strategy in acute hypercapnic respiratory failure due to COPD exacerbation while emphasising close monitoring for early signs of treatment failure,[7] and the 2025 Global Initiative for Chronic Obstructive Lung Disease (GOLD) report similarly recommends NIV as the preferred initial modality and highlights the importance of timely recognition of non-response to avoid delayed intubation.[21] Our findings help to operationalise these recommendations by providing objective early physiological indicators (post-NIV pH and $\Delta PaCO_2$ reduction) that may support such monitoring.

Data from ICU populations demonstrate considerable morbidity and mortality among hypercapnic COPD patients.[22] However, most of these analyses evaluate outcomes after ICU admission rather than evaluating predictors available at presentation to the emergency department. In addition, systematic reviews confirm

Table 2. Comparison of clinical, demographic and laboratory characteristics according to intubation requirement			
Parameters	**No intubation (n=249)**	**Intubation (n=70)**	**p value**
Age, years	70.0±11.1	71.3±12.6	0.403
Hemoglobin, g/dL	13.3 (11.8-15.1)	12.5 (11.2-14.5)	**0.016**
Leukocyte count, /mm^3	10.960 (7.800-13 650)	12 440 (9 100-15 677)	0.067
CRP, mg/L	27.4 (8.1-68.1)	55.0 (23.4-86.7)	**<0.001**
pH (baseline)	7.31 (7.27-7.35)	7.28 (7.24-7.33)	**<0.001**
pH (post-NIV)	7.37 (7.32-7.40)	7.28 (7.22-7.34)	**<0.001**
PaO_2 (baseline), mmHg	62.9 (52.9-72.9)	67.5 (55.0-78.0)	0.090
PaO_2 (post-NIV), mmHg	66.4 (53.9-79.1)	71.25 (55.5-88.9)	**0.039**
HCO_3- (baseline), mmol/L	30.72±4.99	30.33±6.65	0.592
HCO_3- (post-NIV), mmol/L	31.39±6.09	32.64±12.73	0.250
$PaCO_2$ (baseline), mmHg	65.7 (56.4-72.8)	69.8 (60.1-90.5)	**0.022**
$PaCO_2$ (post-NIV), mmHg	58.9 (50.2-65.7)	72.8 (60.1-95.4)	**<0.001**
Δ$PaCO_2$ (post-baseline), mmHg	-7.4 (-14.1 to -3.5)	-0.8 (-11.2 to 13.4)	**<0.001**
Length of hospital stay, days	9 (5-13)	15.5 (9-23)	**<0.001**
Male sex, n (%)	169 (67.9)	53 (75.7)	0.210
Smoking status, n (%)			
Non-smoker	41 (16.5)	9 (12.9)	**0.015**
Ex-smoker	105 (42.2)	43 (61.4)	
Current smoker	103 (41.4)	18 (25.7)	
Hypertension, n (%)	148 (59.4)	47 (67.1)	0.241
Diabetes mellitus, n (%)	57 (22.9)	20 (28.6)	0.557
Cardiovascular disease, n (%)	83 (33.3)	26 (37.1)	0.553
History of cerebrovascular disease, n (%)	8 (3.2)	4 (5.7)	0.336
Previous ICU admission, n (%)	79 (31.7)	52 (74.3)	**<0.001**
History of intubation, n (%)	15 (6.0)	28 (40.0)	**<0.001**
Home oxygen therapy, n (%)	166 (66.7)	56 (80.0)	**0.032**
Home BPAP use, n (%)	90 (36.1)	28 (40.0)	0.562
Hospitalization within previous year, n (%)	214 (85.9)	69 (98.6)	**0.003**

Data are expressed as mean ± standard deviation or median (25th-75th interquartile range) for continuous variables, and as number (percentage) for categorical variables. p values were calculated using appropriate parametric or non-parametric tests based on data distribution. Statistically significant p values are shown in bold. BPAP: Bilevel positive airway pressure, CRP: C-reactive protein, HCO_3-: Bicarbonate, ICU: Intensive care unit, NIV: Non-invasive ventilation, $PaCO_2$: Arterial partial pressure of carbon dioxide, PaO_2: Arterial partial pressure of oxygen, Δ$PaCO_2$: Change in arterial carbon dioxide level calculated as post-NIV, $PaCO_2$ minus baseline $PaCO_2$

that although NIV reduces mortality and intubation rates overall, failure rates remain clinically significant, underscoring the need for early predictive tools.[23] Case-based and management-oriented discussions further emphasize the difficulty clinicians face in determining the optimal timing of intubation for hypercapnic respiratory failure.[24] Our model directly addresses this decision point by providing objective early indicators of potential NIV failure.

With increasing availability of advanced analytical methods, machine-learning-based prediction models have recently been introduced. Jiang et al. developed and validated a machine-learning model to predict the need for invasive mechanical ventilation in AECOPD patients with respiratory failure, demonstrating high discriminative performance.[25] While such approaches may improve predictive accuracy, they often require complex computational infrastructure and may limit bedside interpretability. Notably, our regression-based model achieved comparable discrimination while remaining transparent and implementable through a simple Excel-based calculator, potentially enhancing clinical applicability.

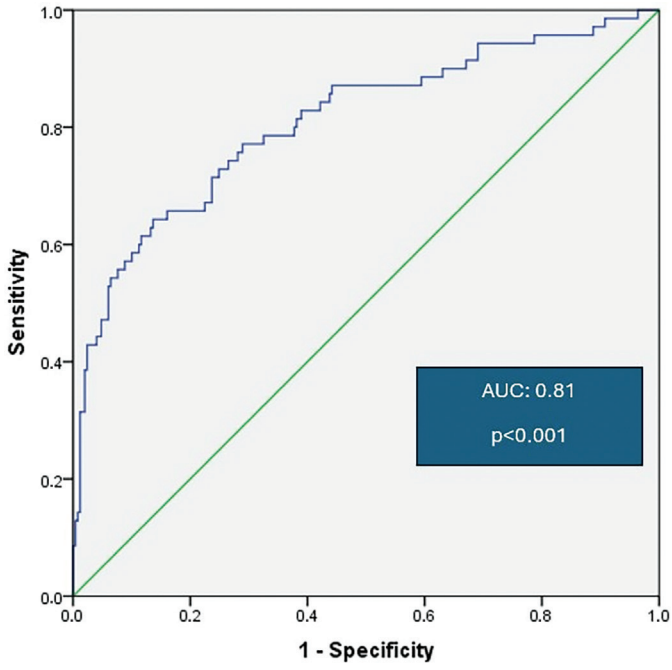


Figure 2. Receiver operating characteristic curve of the regression-based model predicting in-hospital endotracheal intubation in patients with acute hypercapnic respiratory failure due to chronic obstructive pulmonary disease exacerbation treated with non-invasive ventilation in the emergency department. The model showed good discriminative ability with an AUC of 0.810 (95% confidence interval: 0.746-0.873; p<0.001)
AUC: Area under the curve

Study Limitations

This study has several notable strengths. It is based on a multicenter cohort that includes patients from both secondary- and tertiary-level institutions, thereby enhancing the generalizability of the findings across healthcare settings. In contrast to many prior studies focusing predominantly on intensive care populations, this study evaluates patients in whom NIV was initiated in the emergency department, which represents a critical and underexplored time window for clinical decision-making. Moreover, the incorporation of dynamic changes in arterial blood gases, rather than reliance solely on baseline measurements, allows the assessment of the early physiological response to therapy and provides a more clinically meaningful evaluation of treatment trajectory.

Nevertheless, several limitations should be considered. The retrospective design inherently carries a risk of selection bias and residual confounding. In addition, ICU admission was used as one of the primary outcomes; however, because the decision to admit a patient to intensive care may be influenced by external factors such as ICU bed availability, institutional admission policies, and individual physician preference, it may not represent a fully objective clinical endpoint. Endotracheal intubation and in-hospital mortality are less susceptible to these influences and may be regarded as more robust outcome measures. Baseline COPD severity indicators (e.g., GOLD stage, FEV_1, and mMRC score) were not consistently available in this emergency department cohort, which limits full characterisation of disease severity. External validation in independent cohorts was not performed, and formal calibration analyses were not conducted. In addition, long-term outcomes beyond hospital discharge were not assessed, thereby limiting the interpretation to short-term, in-hospital endpoints. These factors should be taken into account when interpreting the predictive performance of the model.

CONCLUSION

Early physiological response to NIV provides important prognostic information in patients with COPD exacerbation complicated by acute hypercapnic respiratory failure. In this multicenter cohort, lower post-treatment pH, insufficient reduction in $PaCO_2$, and a history of prior ICU admission were independently associated with both ICU requirement and in-hospital intubation. The regression-based model incorporating these variables demonstrated good discriminative performance and may support early risk stratification in emergency departments following external validation. External validation in independent populations and confirmation through prospective studies are necessary to further establish its clinical utility.

Ethics

Ethics Committee Approval: This study was approved by the Clinical Research Ethics Committee of İzmir Bakırçay University Faculty of Medicine (approval no: 2702, date: 04.03.2026). The study was conducted in accordance with the principles of the Declaration of Helsinki.

Informed Consent: As this was a retrospective observational study using anonymized data, the requirement for informed consent was waived by the Ethics Committee.

Footnotes

Authorship Contributions

Surgical and Medical Practices: D.S.U., N.E.T., Ö.S.U., M.O.G., E.S.D., H.K., Concept: D.S.U., Ö.S.U., Design: D.S.U., Ö.S.U., Data Collection or Processing: D.S.U., Ö.S.U., H.K., A.M., Analysis or Interpretation: D.S.U., Ö.S.U., A.M., Literature Search: D.S.U., N.E.T., M.O.G., A.M., Writing: D.S.U., M.O.G., A.M.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. World Health Organization. Chronic obstructive pulmonary disease (COPD). Geneva: WHO; 2023. Available from: https://www.who.int/news-room/fact-sheets/detail/chronic-obstructive-pulmonary-disease-(copd)
2. GBD 2019 Chronic Respiratory Diseases Collaborators. Global burden of chronic respiratory diseases and risk factors, 1990-2019: an update from the Global Burden of Disease Study 2019. EClinicalMedicine. 2023;59:101936.
3. Buist AS, McBurnie MA, Vollmer WM, et al. International variation in the prevalence of COPD (the BOLD Study): a population-based prevalence study. Lancet. 2007;370:741-50.
4. Agustí A, Celli BR, Criner GJ, et al. 2025 Global Initiative for Chronic Obstructive Lung Disease 2023 report: GOLD executive summary. Eur Respir J. 2023;61:2300239.
5. Wedzicha JA, Seemungal TA. COPD exacerbations: defining their cause and prevention. Lancet. 2007;370:786-96.
6. Connors AF Jr, Dawson NV, Thomas C, et al. Outcomes following acute exacerbation of severe chronic obstructive lung disease. The SUPPORT investigators (Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments). Am J Respir Crit Care Med. 1996;154:959-67.
7. Rochwerg B, Brochard L, Elliott MW, et al. Official ERS/ATS clinical practice guidelines: noninvasive ventilation for acute respiratory failure. Eur Respir J. 2017;50:1602426.
8. Squadrone E, Frigerio P, Fogliati C, et al. Noninvasive vs invasive ventilation in COPD patients with severe acute respiratory failure deemed to require ventilatory assistance. Intensive Care Med. 2004;30:1303-10.
9. Osadnik CR, Tee VS, Carson-Chahhoud KV, Picot J, Wedzicha JA, Smith BJ. Non-invasive ventilation for the management of acute hypercapnic respiratory failure due to exacerbation of chronic obstructive pulmonary disease. Cochrane Database Syst Rev. 2017;7:CD004104.
10. Davidson AC, Banham S, Elliott M, et al. BTS/ICS guideline for the ventilatory management of acute hypercapnic respiratory failure in adults. Thorax. 2016;71(Suppl 2):ii1-ii35.
11. Demoule A, Girou E, Richard JC, Taille S, Brochard L. Benefits and risks of success or failure of noninvasive ventilation. Intensive Care Med. 2006;32:1756-65.
12. Confalonieri M, Garuti G, Cattaruzza MS, et al. A chart of failure risk for noninvasive ventilation in patients with COPD exacerbation. Eur Respir J. 2005;25:348-55.
13. Ozyilmaz E, Ugurlu AO, Nava S. Timing of noninvasive ventilation failure: causes, risk factors, and potential remedies. BMC Pulm Med. 2014;14:19.
14. Hartley T, Lane ND, Steer J, et al. The noninvasive ventilation outcomes (NIVO) score: prediction of in-hospital mortality in exacerbations of COPD requiring assisted ventilation. Eur Respir J. 2021;58:2004042.
15. Echevarria C, Steer J, Heslop-Marshall K, et al. Validation of the DECAF score to predict hospital mortality in acute exacerbations of COPD. Thorax. 2016;71:133-40.
16. Duan J, Wang S, Liu P, et al. Early prediction of noninvasive ventilation failure in COPD patients: derivation, internal validation, and external validation of a simple risk score. Ann Intensive Care. 2019;9:108.
17. van Gemert JP, Brijker F, Witten MA, Leenen LP. Intubation after noninvasive ventilation failure in chronic obstructive pulmonary disease: associated factors at emergency department presentation. Eur J Emerg Med. 2015;22:49-54.
18. Brochard L, Mancebo J, Wysocki M, et al. Noninvasive ventilation for acute exacerbations of chronic obstructive pulmonary disease. N Engl J Med. 1995;333:817-22.
19. Plant PK, Owen JL, Elliott MW. Early use of non-invasive ventilation for acute exacerbations of chronic obstructive pulmonary disease on general respiratory wards: a multicentre randomised controlled trial. Lancet. 2000;355:1931-5.
20. Chandra D, Stamm JA, Taylor B, et al. Outcomes of noninvasive ventilation for acute exacerbations of chronic obstructive pulmonary disease in the United States, 1998-2008. Am J Respir Crit Care Med. 2012;185:152-9.
21. Global Initiative for Chronic Obstructive Lung Disease. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease: 2025 Report. GOLD; 2024. Available from: https://goldcopd.org/2025-gold-report/
22. Akbaş T, Güneş H. Characteristics and outcomes of patients with chronic obstructive pulmonary disease admitted to the intensive care unit due to acute hypercapnic respiratory failure. Acute Crit Care. 2023;38:49-56.
23. Abualhamael SA, Alasmi AT, Alqurayqiri AF, et al. Role of non-invasive ventilation (NIV) in managing acute exacerbations of chronic obstructive pulmonary disease (COPD): a systematic review. Cureus. 2024;16:e67418.
24. Balkan B, Kaya E, Gokoglu G, Balkan AO, Yilmaz G. Management strategies for hypercapnic respiratory failure in chronic obstructive pulmonary disease (COPD): a case-based discussion. Cureus. 2025;17:e84717.
25. Jiang X, Li J, Ju J, Ding H, Yang S. Construction and validation of a machine learning-based risk prediction model for invasive mechanical ventilation in AECOPD patients complicated with respiratory failure. Int J Med Inform. 2026;208:106244.

DOI: 10.4274/forbes.galenos.2026.43926
Forbes J Med 2026;7:98-106

Original Article

The Relationship Between Wellbeing, Self-Compassion, and Adaptation to Chronic Diseases in Older Adults

Kader MERT[1], Hülya BULUT[2]

[1]İzmir Bakırçay University Faculty of Health Sciences, Department of Nursing, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital, İzmir, Türkiye

Cite as: Mert K, Bulut H. The relationship between wellbeing, self-compassion, and adaptation to chronic diseases in older adults. Forbes J Med. 2026;7:98-106

ABSTRACT

Objective: The aim of this study is to investigate the relationship between well-being, self-compassion, and adaptation to chronic diseases in older adults.

Methods: The descriptive and correlational study was conducted through face-to-face interviews with older adults with chronic disease who were inpatients at a University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital in İzmir between March 2024 and June 2024. A sampling method for known populations was used to select the sample (n=405). A sociodemographic questionnaire, the Warwick-Edinburgh Mental Wellbeing scale, the self-compassion scale, and the Assessment Scale for Adaptation to Chronic Disease were used as instruments for data collection.

Results: The mean age of the patients was 70.69±3.96 years, 58.3% were male. A positive correlation was found between the participants' levels of mental well-being and their adaptation to chronic disease (β=0.389) (p<0.01).

Conclusion: The well-being levels of older adults positively affect adaptation to chronic disease.

Keywords: Adaptation to chronic disease, self-compassion, wellbeing

Received: 17.07.2026
Accepted: 15.08.2026
Epub: 19.08.2026
Publication Date: 28.08.2026

Corresponding Author:
Kader MERT
İzmir Bakırçay University Faculty of Health Sciences, Department of Nursing, İzmir, Türkiye
kader.mert@bakircay.edu.tr
ORCID ID: 0000-0002-5035-7230

INTRODUCTION

Chronic illnesses associated with aging, such as diabetes, hypertension, and cardiovascular disorders, adversely affect older adults' health, particularly in developed countries.[1] The World Health Organization (WHO) states that chronic illnesses are long-term non-communicable diseases that typically progress slowly. WHO data show that 41 million people die of chronic diseases each year worldwide, accounting for 74% of global deaths.[2] TURKSTAT data show that 78.7% of adults over the age of 65 have at least one chronic illness, such as hypertension, diabetes, or heart disease. These illnesses have a serious impact on older adults' activities of daily living and increase their dependency.[3] In this context, well-being and self-compassion can facilitate adaptation to chronic diseases.

For older adults with chronic illnesses who wish to preserve their quality of life, well-being is a crucial concept. The state of well-being is not limited to physical health; it is also determined by how individuals cope with their illness, the extent of the social support they receive, and their overall satisfaction with life.[4] Mental well-being means that the individual is able to cope with stress, form healthy relationships, and contribute to society.[5] Studies show that people with an elevated state of mental well-being have stronger immune systems, engage in healthy social relationships, and live longer.[6] Self-compassion is defined as individuals' ability to treat themselves with kindness and compassion.[7] This notion refers to the individual's acceptance that the difficulties faced in life and the pain and disappointments one encounters are a natural part of life and not a cause for merciless self-criticism.[8] Self-kindness is the ability to treat oneself with compassion when facing challenges; a sense of common humanity involves understanding that others may be experiencing similar difficulties.[9,10] While self-compassion increases an individual's ability to be kind to oneself, it also reduces self-

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

criticism and reminds the person that they are not alone. This helps maintain an emotional equilibrium resulting from accepting that negativities are, by nature, part of the human experience.[11]

As the aging process progresses, the number of chronic diseases increases, and these conditions affect older adults' physical, mental, and social well-being. The aim of this study is to investigate the relationship among well-being, self-compassion, and adaptation to chronic diseases in older adults.

Research questions

1. Is there a relationship between older adults' well-being and their adaptation to chronic diseases?

2. Is there a relationship between older adults' self-compassion and their adaptation to chronic diseases?

METHODS

This is a descriptive and correlational study. This research was conducted from March to June 2024 among patients 65 and older who were inpatients at University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital for pulmonary diseases.

Sample

The study population consisted of inpatients at University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and Surgery Training and Research Hospital for pulmonary diseases who were deemed competent to provide informed consent and were conscious, cooperative, and willing to participate in the research. The number of patients included in the sample was calculated using the formula for determining sample size for a known population. Using the formula n=N.t2 pq/d2(N–1) + t2pq, the sample size was computed to be 405 participants. Thus, the study sample comprised 405 inpatients aged 65 and older who volunteered for the research and had at least one chronic illness. According to the post-hoc power analysis conducted after the study, the effect size (ρ) was found to be 0.329.

Significance level (α) 0.05, and the study's power (1-β) 100%.

Inclusion Criteria: Being age 65 or older, having at least one chronic illness, and consenting to participate in the study.

Exclusion Criteria: Since the instruments used in the study relied on self-reporting, individuals who could not communicate verbally, were not fluent in Turkish, had a mental disability that could hinder comprehension of or response to the questions, or were diagnosed with Alzheimer's disease were excluded from the study.

Data Collection Instruments

The descriptive data questionnaire prepared by the researchers based on the literature consisted of nineteen items.[1,12] The items queried the patients' income status, their smoking and alcohol consumption status, and perceived health status.

The Warwick-Edinburgh Mental Wellbeing scale (WEMWBS) was used to measure the participants' level of mental wellbeing. Developed by Tennant et al.[13], this measure was adapted to Turkish by Keldal.[5] Higher scores indicate better mental health. Cronbach's alpha coefficient for the WEMWBS is 0.89.

The self-compassion scale was developed by Neff[7] (2003a) to measure self-compassion. The scale is a Likert-type instrument that allows individuals to provide information about themselves. The self-compassion scale consists of 26 items.[7] The scale adapted by Deniz et al.[14] reported a single-factor structure. In the Turkish adaptation, item analysis indicated an internal consistency reliability coefficient of 0.89.

Assessment Scale for Adaptation to Chronic Diseases (ASACD) was developed by Kacaroğlu Vicdan and Birgili[15] to assess individuals' adaptation to chronic illnesses; validity and reliability testing was carried out. In the validity and reliability study, Cronbach's alpha reliability coefficient was 0.83. The respondent's adaptation to chronic ailments increases as the score rises.

Data Collection

The researchers collected data while the patients were receiving treatment at the clinics and in the patients' rooms.

Data Evaluation

Data were presented as numerical values, percentage distributions, means, and standard deviations. The results were evaluated using 95% confidence intervals and a significance level of $p<0.05$.

The suitability of the data for a normal distribution was determined using a normality test. Statistical analyses included the t-test, chi-squared test, correlation coefficient (r), Independent samples t-test, one-way ANOVA, Kruskal-Wallis test, Mann-Whitney U test, and Spearman's correlation analysis. Tukey's HSD post-hoc test was employed to ascertain differences among several groups.

Ethics Approval

Ethical approval for the study was obtained from the İzmir Bakırçay University Ethics Committee (dated: 21.02.2024, approval number: 1463) and from the University of Health Sciences Türkiye, İzmir Dr. Suat Seren Chest Diseases and

Surgery Training and Research Hospital where the study was conducted (dated: 28.02.2024, approval number: 4/4). The participants were informed in detail about the objective of the study and the confidentiality of the information, after which their written and verbal consent was obtained.

RESULTS

The average age in the study is 65 years and over (70.69±3.96); 87.7% were between the ages 65-75, 58.3% were males, 72.1% were primary school graduates, 83.2% were married, 71.4% were overweight (body mass index =25-29.9), 66.2% reported income less than their expenditure (Table 1).

A significant difference ($p<0.05$) was observed between participants' ages and the ASACD physiological subscale; however, no difference was detected between participants' ages and the ASACD overall mean score ($p>0.05$). The study indicated a difference ($p<0.05$) in the ASACD social and spiritual subscales by participants' civil status; however, no difference was observed in the overall mean score of the ASACD by civil status ($p>0.05$). The study revealed a significant difference ($p<0.05$) in the ASACD spiritual

Table 1. Sociodemographic health and disease related characteristics of participants (n=405)

Sociodemographic health and disease related characteristics		**(n/%)**
Age	65-75 years 76-85 years	355 (87.7) 50 (12.3)
Age (year) mean (SD) 70.69 (3.96)		
Gender	Woman Man	169 (41.7) 236 (58.3)
Education	Primary education High school and above	292 (72.1) 113 (27.9)
Marital status	Married Single	337 (83.2) 68 (16.8)
Income status	Income less than expenses Income is equal to expenses Income is more than expenses	268 (66.2) 122 (30.1) 15 (3.7)
Smoking status	Using Not using Quit smoking	44 (10.9) 52 (12.8) 309 (76.3)
Current alcohol using status	Using Not using Quit drinking alcohol	4 (1.0) 250 (61.7) 151 (37.3)
BMI	Normal weight (19.03-24.99) Overweight (25-29.90) Obese (30 and above)	51 (12.6) 289 (71.4) 65 (16.0)
Perceived health status	Good Medium Bad	19 (4.7) 247 (61.0) 139 (34.3)
***Chronic disease**	Cardiovascular disease Cancer Chronic respiratory disease Diabetes mellitus Thyroid	313 (77.3) 247 (61.0) 244 (60.2) 209 (51.6) 104 (25.7)
Hospitalization in the last 6 month	Yes No	285 (70.4) 120 (29.6)
Status of meeting self-care needs	No, I can't Sometimes Yes, I can	101 (24.9) 263 (64.9) 41 (10.1)
The situation where disease prevents planned work	Yes No	150 (37.0) 255 (63.0)
Difficulty coping with the disease	Yes No	248 (61.2) 157 (38.8)

*More than one option is marked. SD: Standard deviation, BMI: Body mass index

Table 2. The ASACD mean scores according to sociodemographic characteristics (n=405)						
	n (%)	Physiological sub-scale	Psychological sub-scale	Social sub-scale	Spiritual sub-scale	Total
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Age						
65-75	355 (87.7)	7.31 (2.16)	30.21 (5.85)	10.57 (2.38)	6.59 (1.94)	54.69 (8.15)
76-85	55 (12.3)	7.94 (1.96)	30.16 (4.40)	10.76 (2.01)	6.58 (1.92)	55.44 (6.40)
*Statistics		t=-1.928 p=0.055	t=0.063 p=0.950	t=-0.515 p=0.607	t=0.039 p=0.969	t=-0.616 p=0.538
Gender						
Woman	169 (41.7)	7.53 (2.17)	30.71 (5.86)	10.38 (2.69)	6.52 (2.05)	55.15 (8.45)
Man	236 (58.3)	7.29 (2.13)	29.84 (5.54)	10.75 (2.05)	6.63 (1.86)	54.52 (7.58)
*Statistics		t=1.108 p=0.269	t=1.507 p=0.133	t=-1.500 p=0.135	t=-0.557 p=0.578	t=0.778 p=0.437
Education						
Primary school	292 (72.1)	7.34 (2.13)	29.95 (5.43)	10.61 (2.36)	6.62 (1.88)	54.53 (7.67)
High school and above	113 (27.9)	7.51 (2.18)	30.86 (6.26)	10.56 (2.29)	6.49 (2.07)	55.44 (8.63)
*Statistics		t=-0.702 p=0.483	t=-1.365 p=0.174	t=0.179 p=0.858	t=0.609 p=0.543	t=-1.027 p=0.305
Civil status						
Married	337 (83.2)	7.33 (2.12)	30.16 (5.63)	10.79 (2.21)	6.71 (1.84)	55.00 (8.08)
Single	68 (16.8)	7.66 (2.25)	30.42 (5.99)	9.64 (2.71)	5.97 (2.28)	53.70 (7.24)
*Statistics		t=-1.131 p=0.259	t=0.307 p=0.728	t=3.266 **p=0.002**	t=2.526 **p=0.013**	t=1.233 p=0.218
Income status						
Income less than expenses	268 (66.2)	7.30 (2.15)	30.19 (5.50)	10.66 (2.41)	6.72 (1.88)	54.88 (7.87)
Income equal to expensest	122 (30.1)	7.50 (2.15)	29.90 (6.04)	10.51 (2.20)	6.44 (2.02)	54.36 (8.00)
Income more than expenses	15 (3.7)	8.00 (2.00)	33.00 (5.50)	10.20 (2.24)	5.46 (1.95)	56.66 (9.21)
**Statistics		KW =1.944 p=0.378	KW =4.065 p=0.131	KW =1.981 p=0.371	KW =6.020 **p=0.049 1-3; 3-1**	KW =0.978 p=0.613
BMI						
Normal (19.03-24.99)	51 (12.6)	7.37 (2.21)	29.68 (5.48)	10.39 (2.17)	6.4 (1.81)	53.90 (7.90)
Overweight (25-29.90)	289 (71.4)	7.19 (2.11)	30.30 (5.84)	10.49 (2.39)	6.57 (2.03)	54.57 (8.06)
Obese (30 and above)	65 (16.0)	8.29 (2.09)	30.18 (5.20)	11.21 (2.15)	6.75 (1.59)	56.44 (7.36)
***Statistics		F=7.128 **p=0.001 2-3; 3-2**	F=0.256 p=0.775	F=2.734 p=0.066	F=0.367 p=0.693	F=1.840 p=0.160
Perceived health status						
Good	19 (4.7)	8.26 (2.53)	30.78 (7.20)	10.00 (2.88)	6.05 (2.41)	55.10 (10.57)
Medium	247 (61.0)	7.47 (2.14)	30.67 (5.71)	10.68 (2.22)	6.66 (1.91)	55.49 (8.07)
Bad	139 (34.3)	7.12 (2.08)	29.29 (5.33)	10.53 (2.47)	6.53 (1.92)	53.48 (7.19)
**Statistics		KW =3.762 p=0.152	KW =5.273 p=0.072	KW =0.815 p=0.665	KW =1.120 p=0.571	KW =6.625 **p=0.036 2-3; 3-2**
Total 405 100		7.39 (2.15)	30.00 (5.69)	10.60 (2.34)	6.59 (1.94)	54.79 (7.95)

*t= Independent sample t-test, **Kruskall Wallis, ***F= One-way ANOVA. SD: Standard deviation, BMI: Body mass index, ASACD: Assessment Scale for Adaptation to Chronic Diseases

subscale across individuals' income levels; however, there was no statistically significant difference between economic status and the ASACD total mean score (p>0.05). A significant difference was observed between the participants' perceived health status and the ASACD overall and subscale mean scores (p<0.05). It was observed that those who perceived their health to be "average" had significantly higher mean scores in their adaptation to their illness (KW =6.625, p=0.036) (Table 2).

A significant difference was observed when comparing the participants' hospital admissions in the last 6 months with their social subscale and total ASACD mean scores (p<0.05). A difference in knowledge about the illness was observed when comparing the physiological and psychological subscales with the overall ASACD mean score (p=0.003); in the advanced analysis, variances were homogeneous (p>0.05), and therefore the Tukey honestly significant difference (HSD) test was applied.

Table 3. Assessment Scale for Adaptation to Chronic Disease (ASACD) mean scores according to disease related characteristics and correlation of the WEMWBS and self compassion scale with the ASACD (n=405)

	n (%)	Physiological sub-scale	Psychological sub-scale	Social sub-scale	Spiritual sub-scale	Total
		Mean (SD) 7.39 (2.15)	**Mean (SD)** 30.00 (5.69)	**Mean (SD)** 10.60 (2.34)	**Mean (SD)** 6.59 (1.94)	**Mean (SD)** 54.79 (7.95)
Hospitalization in the last 6 month						
Yes	285 (70.4)	7.18 (2.04)	29.85 (5.50)	10.38 (2.27)	6.47 (1.95)	53.89 (7.47)
No	120 (29.6)	7.88 (2.32)	31.04 (6.04)	11.11 (2.43)	6.87 (1.87)	56.91 (8.65)
*Statistics		t=-3.008 **p=0.003**	t=-1.921 p=0.055	t=-2.904 **p=0.004**	t=-1.923 p=0.055	t=-3.335 **p=0.001**
Level of knowledge about disease						
Insufficient	113 (27.9)	6.97 (2.11)	29.07 (4.87)	10.55 (2.44)	6.60 (1.86)	53.20 (6.71)
Moderately sufficient	274 (67.7)	7.48 (2.14)	30.44 (5.84)	10.64 (2.26)	6.64 (1.90)	55.21 (8.21)
Very sufficient	18 (4.4)	8.66 (2.02)	33.77 (6.34)	10.22 (2.88)	5.66 (2.65)	58.33 (9.51)
**Statistics		KW =9.278 **p=0.010**	KW =13.403 **p=0.001 1,2<3**	KW =0.313 p=0.855	KW =2.116 p=0.347	KW =11.328 **p=0.003**
Status of meeting self-care needs						
No, I can't	101 (24.9)	7.12 (2.00)	29.30 (5.30)	10.62 (2.43)	6.53 (1.94)	53.59 (7.17)
Sometimes	263 (64.9)	7.22 (2.04)	29.98 (5.47)	10.61 (2.21)	6.66 (1.86)	54.48 (7.38)
Yes, I can	41 (10.1)	9.12 (2.41)	33.85 (6.66)	10.46 (2.95)	6.26 (2.39)	59.70 (11.17)
***Statistics		F=15.897 **p=0.000 1,2<3**	F=10.341 **p=0.000 1,2<3**	F=0.078 p=0.925	F=0.782 p=0.458	F=9.558 **p=0.000 1,2<3**
The situation where disease prevents planned work						
Yes	150 (37.0)	7.22 (2.03)	29.26 (5.35)	10.45 (2.45)	6.48 (1.89)	53.42 (7.42)
No	255 (63.0)	7.42 (2.16)	30.57 (5.77)	10.67 (2.27)	6.63 (1.96)	55.30 (8.11)
*Statistics		t=-0.907 p=0.365	t=-2.252 **p=0.025**	t=-0.899 p=0.369	t=-0.718 p=0.473	t=-2.301 **p=0.022**
Difficulties coping with the disease						
Yes	248 (61.2)	7.18 (2.09)	29.43 (5.48)	10.57 (2.41)	6.49 (1.97)	53.68 (7.53)
No	157 (38.8)	7.72 (2.20)	31.42 (5.81)	10.64 (2.23)	6.73 (1.87)	56.52 (8.31)
*Statistics		t=-2.498 **p=0.013**	t=-3.459 p=0.001	t=-0.295 p=0.768	t=-1.228 p=0.220	t=-3.549 **p=0.000**
Warwick-Edinburgh mental well being scale		r=0.370 **p=0.000**	r=0.354 **p=0.000**	r=0.050 p=0.319	r=0.108 **p=0.030**	r=0.394 **p=0.000**
Self-compassion scale		r=0.099 **p=0.047**	r=0.055 p=0.271	r=0.146 p=0.003	r=0.049 p=0.329	r=0.121 **p=0.015**

*t= Independent sample t-test, **x^2= Kruskall Wallis,***F= One-way ANOVA, rho: Spearman correlation, SD: Standard deviation, WEMWBS: Warwick-Edinburgh Mental Wellbeing scale

The difference was attributable to the participants with low and moderate levels of knowledge. Those who had a "very sufficient" level of knowledge about their illness had significantly higher mean adaptation scores than those with other levels of knowledge. A difference was found between the physiological and psychological subscale scores for meeting self-care needs and the total ASACD mean score ($p<0.01$). In the post-hoc analysis, the variances were homogeneous ($p>0.05$), and the Tukey HSD test was applied (Table 3).

Multiple linear regression analysis revealed that WEMWBS scores accounted for 15.1% of ASACD (adjusted R^2 =0.151; $F=37.064$, $p<0.001$). A unit increase in the wellbeing scale was found to cause a 0.363-unit increase on the ASACD scale ($B=0.363$, $p<0.001$). No statistically significant relationship was found between the self-compassion scale scores and the ASACD ($p>0.05$) (Table 4).

DISCUSSION

Owing to increasing life expectancy in contemporary populations, the incidence of chronic diseases has risen.[16] This study was conducted with a group of individuals with different chronic diseases rather than with individuals with a single disease; when the individuals' adaptation to their illness was assessed, it was found to be at a moderate level. In the study by Abdan[17] individuals diagnosed with chronic illness exhibited an adaptation mean score of (66.38±7.83). Diagnoses among the participants in this study were cardiovascular disease (77.3%), cancer (61%), chronic airway diseases (60.2%), diabetes mellitus (51.6%), and thyroid disease (25.7%). There are similar studies in the literature. In the study by Bilgiç and Pehlivan[16] conducted with a group of individuals with different chronic diseases instead of with a group with only a particular disease, it was found that the individuals' adaptation to their illness was at above moderate. Some studies examining hypertensive patients' adaptation to drug therapy found high levels of adaptation.[18,19] Another study, however, noted that adaptation to therapy was low.[20] Researchers working with individuals with MS found low levels of adaptation to treatment.[21] A study of individuals with diabetes reported that psychosocial adaptation to the illness was moderate.[22]

No statistically significant correlation was identified between participants' age and their ability to adapt to their illnesses. Özdemir et al.[23] similar to Abdan[17] no significant association was discovered between age groups and adaption to chronic illness. This result shows that adaptation to chronic disease is not simply tied to age, but other factors can also play an important role in this process. In the study by Cezik[24], participants with hypertension had significantly higher ASACD scores. According to the study data, no significant relationship was found between the participants' gender and their adaptation to chronic disease ($p=0.437$). The results reported in the literature are similar to ours. For example, Abdan[17] and Cezik[24] did not detect a significant relationship between chronic disease and gender in their studies, concluding that gender may not be a determining factor in the adaptation process. The results of Abdan's study[17] support this finding. The lack of a significant association between participants' gender and adherence to antihypertensive therapy in the study conducted by Özdemir et al.[23] on geriatric hypertensive patients reflects the general trend observed in the literature. On the other hand, Gupta et al.[25] reported in their study conducted in the Czech Republic and in the UK that women show 55% greater maladaptation than men, which suggests that in some cases, gender can have an impact on the adaptation process. In contrast, Shea et al.[26] and Ross et al.[27] and other researchers have shown that women's adaptation to treatment is better than men's. These contradictory findings suggest that the impact of gender on adaptation to chronic disease is associated not only with biological factors but also with socioeconomic conditions, cultural norms, regional differences, and other environmental factors. When considering gender as a determining factor in adaptation to chronic disease, social and cultural factors should also be evaluated. Various findings in the literature stress that considering gender as one of many dynamics influencing adaptation would be a better approach. Health statistics indicate differences in the distribution of chronic illnesses by gender and type.[28] In this study, the number of males with chronic disease (58.3%, n=236) was greater than that of females, which is contrary to what is reported in the literature.

Table 4. According to multiple linear regression analysis the relationship between in individuals' well being and self-compassion scores and adaptation to chronic disease

Model assessment scale for adaptation to chronic disease	**B**	**95.0% CI for B**	**β**	**t**	**p**
Warwick-Edinburgh Mental Wellbeing scale	0.363	0.276, 0.450	0.389	8.195	**<0.001**
Self-compassion scale	0.011	-0.043, 0.064	0.018	0.388	0.698
Adjusted R^2: 0.151; F=37.064; p<0.001					

B: Unstandardized coefficients, CI: Confidence interval, β: Standardized regression coefficient, R^2: Coefficient of determination

According to these results, no statistically significant difference was found between participants' civil status and their ASACD scores (p=0.218). Findings in the literature support this result. Bilgiç and Pehlivan[16] and Abdan[17] report that they found no significant relationship between adaptation to chronic disease and civil status, while Özdemir et al.[23] similarly reveal in their study with older adults that civil status does not make a significant impact on adaptation to drug therapy. These studies suggest that civil status alone may not predict the degree of adaptation to chronic disease. On the other hand, some studies report that civil status can affect adaptation to chronic disease. For example, it was found in Cezik's[24] study that divorced/widowed individuals had lower adaptation scores while Cooper et al.[29] reported in a study conducted in 11 countries that older unmarried adults had lower scores related to their adaptation to drug therapy.

This study found no significant difference in ASACD scores by educational status (p=0.305). Erdoğan Yüce and Muz[30], however, did find a relationship between education and adaptation to chronic disease. Individuals with lower educational attainment demonstrated poorer adaptation to chronic disease. Other studies have also reported that adaptation to illness increases with higher levels of patient education.[16,17,24,31-35]

The relationship between chronic illnesses and income status has been explored in many studies. In particular, the number of chronic diseases is higher in societies with lower socioeconomic levels. It must be said, however, that this does not hold true for all chronic diseases: hypertension, breast cancer, and similar chronic illnesses are reported to be much more prevalent in societies where socioeconomic conditions are better.[36]

This study revealed a statistically significant correlation between participants' income levels and the ASACD spiritual subscale (p<0.05); however, no statistically significant difference was observed between participants' economic status and the total mean score of the ASACD (p>0.05). This suggests that, despite the psychological impact of income status on the ASACD, the overall ASACD score did not differ noticeably. Differing from our results in the literature, the studies on the relationship between income and chronic illness of Abdan[17], Cezik[24] Erdoğan Yüce and Muz[30], and Callaghan[37] indicated a significant association between income and adaptation to chronic disease. The WHO reports that low socioeconomic status is associated with poor adaptation to chronic disease.[38] Especially if the chronic illness has led to lifestyle changes and dependence on others, the illness may be more difficult to accept, and adaptation may be more difficult.[39]

A significant difference in ASACD overall and subscale mean scores was found between study participants who were hospitalized in the last 6 months and those who were not (p=0.001). Review of the participants' hospitalization histories showed that 70.4% had been admitted for a chronic disease in the previous 6 months, and 51.1% had been hospitalized at least once during that period. Since chronic illnesses inherently require longer hospital stays, they account for a significant portion of health service costs and job loss.[40]

Multiple linear regression analysis revealed that WEMWBS scores accounted for 15.1% of ASACD (adjusted R^2 =0.151; F=37.064, p<0.001). Since the variables measured in the model in this study are concepts specific to the social sciences and psychology, R-squared values as low as 0.10 to 0.30 are often considered acceptable.[41] Furthermore, regression analysis revealed no statistically significant relationship between self-compassion scale scores and the ASACD. This may be due to the inclusion of sociodemographic and clinical characteristics in the regression analysis.

Study Limitations

Limitations of the study included its conduct in a single facility and its cross-sectional design, conducted during a defined period.

CONCLUSION

A low, positive relationship was detected between the overall ASACD and WEMWBS. There is no significant association between older patients' self-compassion and their adaptation to chronic disease. There is a significant association between their levels of mental well-being and their adaptation to chronic disease. Wellbeing is associated with adaptation to chronic disease. In an individual's later years, a higher sense of mental well-being is conducive to better adaptation to chronic disease. Nurses play an important role in evaluating the well-being of older adults and in planning interventions to increase their well-being, thereby enhancing adaptation to chronic disease. It is recommended that well-being levels be examined in nursing research aimed at improving adaptation to chronic diseases in older adults, and that psychological nursing interventions be planned to enhance well-being.

Ethics

Ethics Committee Approval: Ethical approval for the study was obtained from the İzmir Bakırçay University Ethics Committee (dated: 21.02.2024, approval number: 1463) and from the hospital where the study was conducted (dated: 28.02.2024, approval number: 4/4).

Informed Consent: The participants were informed in detail about the objective of the study and the confidentiality of the information, after which their written and verbal consent was obtained.

Footnotes

Authorship Contributions

Concept: K.M., H.B., Design: K.M., H.B., Data Collection or Processing: H.B., Analysis or Interpretation: K.M., H.B., Literature Search: K.M., H.B., Writing: K.M., H.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Farzad M, MitraMoodi, Sharifi F, et al. Prevalence and factors connected with chronic diseases in the elderly residents of Birjand: a community - based study in Birjand, South Khorasan Province, Iran. J Diabetes Metab Disord. 2021;20:1655-62.
2. World Health Organization (WHO). Non-communicable diseases. 2022. Available from https://www.who.int/news-room/fact-sheets/detail/noncommunicable-diseases.
3. Turkish Statistical Institite (TURKSTAT). Türkiye'de yaşlı profili araştırması 2023. Available from https://veriportali.tuik.gov.tr/tr/press/53809/metadata
4. Anaby D, Miller WC, Jarus T, Eng JJ, Noreau L. Participation and well-being among older adults living with chronic conditions. Soc Indic Res. 2011;100:171-83.
5. Keldal, G. Turkish version of the Warwick-Edinburgh mental well-being scale: a validity and reliability study. J Happiness Well-Being. 2015;3:103-15.
6. Keyes CL, Dhingra SS, Simoes, EJ. Change in level of positive mental health as a predictor of future risk of mental illness. Am J Public Health. 2010;100:2366-71.
7. Neff KD. Self-compassion: an alternative conceptualization of a healthy attitude toward oneself. Self Identity. 2003;2:85-101.
8. Neff KD. The development and validation of a scale to measure self-compassion. Self Identity. 2003;2:223-50.
9. Neff KD. Does self-compassion entail reduced self-judgment, isolation, and over-identification? A response to Muris, Otgaar, and Petrocchi. Mindfulness. 2016;7:791-7.
10. Neff KD. Commentary on Muris and Otgaar Let the empirical evidence speak on the self-compassion scale. Mindfulness. 2020;11:1900-9.
11. Özyeşil Z. Self-compassion and mindfulness. Ankara: Maya Akademi Publishing; 2011. p. 26-34.
12. Kaçan H, Değer VB. Investigation of anxiety and coping levels of hospitalized elderly and chronically diseased individuals and affecting variables. Dicle Med J. 2024;50:117-27.
13. Tennant R, Hiller L, Fishwick R, et al. The Warwick-Edinburgh mental well-being scale (WEMWBS): Development and UK validation. Health Qual Life Outcomes. 2007;5:1-13.
14. Deniz M, Kesici S, Sümer AS. The validity and reliability of the Turkish version of the self-compassion scale. Soc Behav Pers. 2008;36:1151-60.
15. Kacaroğlu Vicdan A, Birgili F. The validity and reliability study for developing an assessment scale for adaptation to chronic diseases. Journal of Current Researches on Health Sector. 2018;8:135-44.
16. Bilgiç Ş, Pehlivan E. The relationship of adaptation to the disease and quality of life of individuals with chronic disease. Journal of Samsun Health Science. 2023;8:63-76.
17. Abdan Ç. Depression and family role change in individuals diagnosed with chronic disease and their relatives. Karabük University Graduate Education Institute, Department of Social Work, Master's Thesis. Karabük, Türkiye; 2023.
18. Aşiret GD, Okatan C. Determination of the relationship between drug compliance levels and spiritual wellbeing of hypertension patients. Turk J Card Nurs. 2019;10:122-8.
19. Teke N, Arslan S. Hypertensive individuals's medication adherence self efficacy level in rural population and determination of the affeting factors. DEUHFED. 2018;11:120-8.
20. Gürdoğan M, Gürdoğan E. Treatment compliance and related factors in Hypertension patients. MN Cardiology. 2019;26:147-53.
21. Erbay Ö, Yeşilbalkan ÖU, Yüceyar A. Factors that affect adherence of disease- modifying theraphy in patients with multiple sclerosis. DEUHFED. 2018;11:164-72.
22. Kaymaz TT, Akdemir N. Psychosocial adjustment to disease in ındividuals with diabetes. J Psy Nurs. 2016;7:61-7.
23. Özdemir Ö, Akyüz A, Doruk H. Compliance with drug treatment in geriatric patients with hypertension. Med J Bakirkoy. 2016;12:195-201.
24. Cezik, ES. Evaluation of chronic disease adaptation of elderly patients diagnosed with hypertension in Edirne Center. Trakya University, Faculty of Medicine department of Family Medicine. Specialization Thesis. Edirne, Türkiye. 2020.
25. Gupta P, Patel P, Štrauch, B, et al. Risk factors for nonadherence to antihypertensive treatment. Hypertension. 2017;69:1113-20.
26. Shea S, Misra D, Ehrlich MH, Field L, Francis CK. Correlates of nonadherence to hypertension treatment in an inner-city minority population. Am J Public Health. 1992;82:1607-12.
27. Ross S, Walker A, MacLeod MJ. Patient compliance in hypertension: role of illness perceptions and treatment beliefs. J Hum Hypertens. 2004;18:607-13.
28. Turkish Statistical Institute (TURKSTAT) 2023. https://data.tuik.gov.tr/Bulten/Index?p=Turkiye-Saglik-Arastirmasi-2022-49747, Publish date: 01 June 2023, Number: 49747 Accessed June 2, 2026.
29. Cooper C, Carpenter I, Katona C, et al. The AdHOC study of older adults' adherence to medication in 11 countries. Am J Geriatr Psychiatry. 2005;13:1067-76.
30. Erdoğan Yüce G, Muz G. Investigation of the relationship betwen health literacy level and adjustment to chronic disease in individulas with chronic disease: a cross-sectional study. Gumushane Univ J Health Sci. 2023;12:115-23.
31. Aslan H, Çetkin T, Dikmen RD. Effects of chronic illness adaptation on the healthy lifestyle behaviours of patients in internal services. Cukurova Med J. 2021;46:1140-9.
32. Karatepe H, Atik D, Yuce, UO. Adaptation with the chronic disease and expectations from nurses. Erciyes Med J. 2020;42:18-25.
33. Hsieh PL, Yang FC, Hu YF, et al. Continuity of care and the quality of life among patients with type 2 diabetes mellitus: a cross-sectional study in Taiwan. Healthcare (Basel). 2020;8:486.
34. Erci B, Elibol M, Aktürk Ü. Evaluation of correlation between adherence to treatment of hypertansive patients and their life quality. Florence Nightingale J Nurs. 2018;26:79-92.

35. Tokem Y, Taşçı E, Yılmaz M. Investigation of disease management of individuals with hypertension at home. Turk J Card Nur. 2013;4:30-40.
36. Lai S, Shen C, Yang X, et al. Socioeconomic inequalities in the prevalence of chronic diseases and preventive care among aged 45 and older in Shaanxi Province, China. BMC Public Health. 2019;19:1460.
37. Callaghan D. Healthy behaviors, self-efficacy, self-care, and basic conditioning factors in older adults. J Community Health Nurs. 2005;22:169-78.
38. World Health Organization and others. Global action plan for the prevention and control of noncommunicable diseases 2013-2020: World Health Organization; 2013:55.
39. Altundağ S. The effect of social support and education at adaptation to disease in children with type 1 diabetes mellitus. Pamukkale Medical Journal. 2018;11:137-44.
40. Anderson E, Durstine JL. Physical activity, exercise, and chronic diseases: a brief review. Sports Med Health Sci. 2019;1:3-10.
41. Chicco D, Warrens MJ, Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. PeerJ Comput Sci. 2021;7:e623.

DOI: 10.4274/forbes.galenos.2026.50133
Forbes J Med 2026;7:115-21

Original Article

Association of Serial Postoperative CASUS Scores With In-Hospital Mortality Among Patients With Prolonged Mechanical Ventilation After Open-Heart Surgery: A Retrospective Cohort Study

Anıl ONUR[1], Osman Sıla AYDIN[1], Ahmet Burak TATLI[2], Cihan Sedat AYTÜNÜR[1], Nuri Burkay SOYLU[1], Mesut ÖZTÜRK[1], Ahmet Kağan AS[2], Ümran KARACA[1], Tuğba ONUR[1]

[1]University of Health Sciences Türkiye, Bursa Yüksek İhtisas Training and Research Hospital, Clinic of Anesthesiology and Reanimation, Bursa, Türkiye
[2]University of Health Sciences Türkiye, Bursa Yüksek İhtisas Training and Research Hospital, Clinic of Cardiovascular Surgery, Bursa, Türkiye

Cite as: Onur A, Aydın OS, Tatlı AB, et al. Association of serial postoperative CASUS scores with in-hospital mortality among patients with prolonged mechanical ventilation after open-heart surgery: a retrospective cohort study. Forbes J Med. 2026;7:115-21

ABSTRACT

Objective: To evaluate the association between serial cardiac surgery score (CASUS) measurements and in-hospital mortality, and to assess the day-specific discrimination of these measurements among patients requiring prolonged mechanical ventilation after open-heart surgery.

Methods: This single-center retrospective cohort screened 1,255 adults who underwent open-heart surgery with cardiopulmonary bypass (CPB). Of the 118 patients ventilated for more than 24 hours, 55 lacked the data required for the postoperative day 1 CASUS calculation, leaving 63 patients for analysis. CASUS was assessed on postoperative days 1-5 using linear mixed-effects models, receiver operating characteristic analysis, and logistic regression analysis.

Results: Twenty-one patients survived, and 42 died in hospital. CASUS availability decreased from 63 on day 1 to 59 on day 5 because four non-survivors died before subsequent measurements could be obtained. The group-by-day interaction was significant ($p<0.001$). Day 4 and day 5 scores were higher among non-survivors (Welch's t-test, $p=0.011$ and $p<0.001$, respectively). Day 5 CASUS showed moderate discrimination [area under the curve 0.729, 95% confidence interval (CI) 0.602-0.856; $p=0.004$]. At a cut-off >8.5, sensitivity was 55.3% and specificity was 90.5%. Among 59 complete cases, day-5 CASUS was associated with in-hospital mortality after adjustment for age, CPB duration, and EuroSCORE II (adjusted odds ratio 1.245, 95% CI 1.076-1.439; $p=0.003$).

Conclusion: Day 5 CASUS was associated with mortality and showed moderate discrimination in this selected cohort, but it represented a late severity marker rather than an early prognostic tool. External validation is required.

Keywords: Cardiac surgical procedures, respiration, artificial, mortality, risk assessment, intensive care units

Received: 02.07.2026
Accepted: 16.09.2026
Publication Date: 29.09.2026

Corresponding Author:
Anıl ONUR
University of Health Sciences Türkiye, Bursa Yüksek İhtisas Training and Research Hospital, Clinic of Anesthesiology and Reanimation, Bursa, Türkiye
dr.anilonur@gmail.com
ORCID ID: 0000-0002-3957-922X

INTRODUCTION

Cardiopulmonary bypass (CPB) exposes patients undergoing open-heart surgery to several inflammatory stimuli. Contact between blood and the artificial circuit, together with operative trauma, endotoxemia, and ischemia-reperfusion, activates leukocytes and the vascular endothelium; proinflammatory cytokines, reactive oxygen metabolites, and nitric oxide are subsequently released. Systemic inflammation is frequent after cardiac surgery, although its clinical relevance and definition vary.[1] When this response becomes dysregulated, multiple organ dysfunction may follow and postoperative morbidity and mortality may increase.

Definitions of prolonged mechanical ventilation (PMV) differ across studies; in cardiac surgery, PMV is commonly defined as ventilation continuing beyond 24 hours. PMV is a major postoperative complication that is associated with greater in-hospital mortality, longer intensive care unit (ICU) stays, more infections, and higher treatment costs. A recent meta-analysis estimated an overall incidence of 20%, including approximately 12% after isolated coronary artery bypass grafting and confirmed an associated increase in mortality.[2] Reported independent predictors include female sex, preoperative leukocytosis, longer CPB exposure, elevated lactate levels, and reoperation for postoperative bleeding.[3] These consequences have prompted efforts both to validate existing tools and to develop new models

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

for identifying patients at risk; a recent study developed and validated one such model.[4]

Various scoring systems have been developed to support clinical decision-making for cardiac surgery ICU patients. EuroSCORE II is widely used to assess preoperative mortality risk, but it was not designed for daily postoperative stratification, while commonly used postoperative scores such as Sequential Organ Failure Assessment (SOFA) and Acute Physiology and Chronic Health Evaluation (APACHE II) lack calibration specific to cardiac surgery patients. Although EuroSCORE II has shown good discrimination for mortality, it has been found unable to reliably predict risk in high-risk patients (score above 10), highlighting the need for new scoring systems.[5,6]

Cardiac surgery score (CASUS) was designed to allow repeated assessment of postoperative mortality in patients in the cardiac surgical ICU. Compared with other ICU scores, CASUS and SOFA both provided reliable estimates, but CASUS achieved better individual-level prediction and larger daily receiver operating characteristic (ROC) areas, leading investigators to recommend it for this setting.[7] The scored items are: glasgow coma scale, PaO_2/FiO_2, creatinine, bilirubin, renal replacement therapy, lactate, platelet count, pressure-adjusted heart rate, and support with an intra-aortic balloon pump or ventricular assist device. The calculation is straightforward and can be performed online at www.cardiac-icu.org. That study also reported high discrimination (AUC ≥0.90).[7]

In our country, data on the clinical validation of the CASUS score in cardiac surgery are limited. At our center, EuroSCORE II is used to predict postoperative mortality, yet the current literature suggests that CASUS provides more reliable results for daily risk assessment. Validation studies at various centers indicate that CASUS may require local modification and that its additive version offers practical advantages over the logistic version, particularly in settings with low mortality rates.[8] This suggests that the predictive value of CASUS in a high-risk, complicated patient group warrants further investigation and constitutes the primary hypothesis of our study.

PMV identifies a clinically severe postoperative subgroup in which evolving organ dysfunction may be reflected in serial CASUS measurements. However, the performance of CASUS in this restricted population has not been adequately characterized. Therefore, this study aimed to describe CASUS scores during the first five postoperative days and to evaluate their association with, and day-specific discriminatory performance for, in-hospital mortality among adults who required PMV (>24 hours) after open-heart surgery. Mortality was the study outcome and was not an inclusion criterion.

METHODS

Study Design and Patient Population

We assembled this retrospective cohort at a single center. Records of 1,255 adults who underwent open-heart surgery with CPB between June 1, 2024, and December 31, 2025 were screened. Among them, 118 patients required invasive mechanical ventilation for more than 24 hours postoperatively and therefore met the PMV definition. Off-pump and non-cardiac procedures were not part of the eligible source population. Fifty-five patients were excluded because the clinical or laboratory data required to calculate the postoperative day 1 CASUS score were incomplete. The remaining 63 patients met the PMV criterion and were categorized according to in-hospital survival: 21 survivors and 42 non-survivors.

The study was approved by the Medical Sciences Ethics Committee of the University of Health Sciences Bursa Yüksek İhtisas Training and Research Hospital on March 11, 2026 (decision no: 2024-TBEK 2026/03-02). Identifying approval details were submitted separately to the editorial office and were withheld from the blinded manuscript. The study was conducted in accordance with the Declaration of Helsinki. Owing to the retrospective design, the ethics committee waived the requirement for informed consent.

Inclusion and Exclusion Criteria

Patients aged 18 years or older who underwent open-heart surgery with CPB during the study period and developed PMV (>24 hours) were eligible. Off-pump and non-cardiac procedures were not eligible. Among patients meeting the PMV criterion, those with incomplete clinical or laboratory data required for the postoperative day 1 CASUS calculation were excluded. In-hospital mortality was not an inclusion criterion.

Data Collection and Variables

Information was abstracted retrospectively from four institutional sources: the encrypted hospital information system, the ICU monitoring records, the clinical charts, and the laboratory database. Collected demographic measures included age, sex, height, weight, and body mass index. Preoperative variables included diabetes mellitus, hypertension, coronary artery disease, chronic obstructive pulmonary disease, chronic kidney disease, ejection fraction, laboratory results, and EuroSCORE II. Procedure-related variables included operation type, CPB and aortic cross-clamp durations, and IABP or VAD use; postoperative variables were also recorded.

We applied the additive version of CASUS developed for cardiac surgical intensive care. Its 10 inputs include lactate, creatinine, bilirubin, pressure-adjusted heart rate, PaO_2/FiO_2, platelet count, Glasgow Coma Scale, IABP and VAD support, and renal replacement therapy. CASUS was calculated on postoperative days 1-5, when the required measurements were available. Postmortem measurements were structurally unavailable and not imputed.

Endpoints

The primary endpoint was in-hospital mortality. PMV (>24 hours) was the cohort-defining inclusion criterion and was not evaluated as an endpoint. The analyses assessed, within this selected PMV cohort only, the association and discriminatory performance of serial CASUS scores for mortality.

Statistical Analysis

Statistical analyses were performed with SPSS version 25.0 (IBM Corp., Armonk, NY, USA). Normality was assessed with the Shapiro-Wilk test. Continuous variables are presented as mean ± standard deviation or median (interquartile range), as appropriate, and categorical variables as number and percentage. Between-group comparisons were performed using the independent-samples t test or Mann-Whitney U test for continuous variables, and the chi-square test or an exact test for categorical variables. Surgical procedure types were compared using the Fisher-Freeman-Halton exact test. The longitudinal CASUS course was examined using a linear mixed-effects model with a patient-specific random intercept, and with mortality group, postoperative day (as a categorical factor), and their interaction included as fixed effects. The model used all observed measurements and accommodated unbalanced follow-up, but it could not eliminate bias arising from informative dropout due to death. Day-specific observed CASUS scores were compared using Welch's t-test.

Day-specific ROC analyses included patients with an available CASUS measurement for the corresponding day; each AUC was tested against 0.5, and exploratory cut-offs were selected using the Youden index. No adjustment for multiple testing was applied; therefore, day-specific comparisons and ROC analyses were considered exploratory. The postoperative day-5 CASUS score was the predictor of interest in the logistic regression. Age, CPB duration, and EuroSCORE II were included on clinical grounds as adjustment covariates, rather than being selected by univariable p values. Univariable and adjusted odds ratios (ORs) with 95% confidence intervals were reported for the common complete-case day-5 population. Model fit was summarized using Nagelkerke R^2 and the Hosmer-Lemeshow test; the latter was not interpreted as evidence of good calibration. No formal a priori sample-size calculation was performed because all eligible patients in the predefined retrospective period were included. A two-sided p value <0.05 was considered statistically significant. This study was reported in accordance with STROBE guidelines.

RESULTS

Of 1,255 patients screened, 118 developed PMV lasting longer than 24 hours. Fifty-five patients were excluded because the clinical or laboratory data required to calculate the postoperative day 1 CASUS score were incomplete, leaving 63 patients in the final analysis. All 63 patients met the PMV criterion: 21 (33.3%) survived and 42 (66.7%) died during hospitalization.

The mean age was 66.1±5.9 years among survivors and 66.1±9.7 years among non-survivors (p=0.980). No significant differences were found in sex distribution (p=0.270), body mass index (p=0.590), diabetes mellitus (p=0.580), hypertension (p=0.780), chronic obstructive pulmonary disease (p=0.470), chronic kidney disease (p=0.510), ejection fraction (p=0.640), or EuroSCORE II (p=0.220). The cohort included 27 CABG operations, 17 combined operations, 11 valve operations, and 8 aortic operations. The type of surgical procedure did not differ by mortality status (Fisher-Freeman-Halton exact p=0.403; Table 1).

Intraoperative characteristics were similar between survivors and non-survivors (Table 2). CASUS measurements were available for 63, 62, 62, 60, and 59 patients on postoperative days 1-5, respectively. All 21 survivors had measurements through day 5. Among non-survivors, one patient died on postoperative day 1, two on day 3, and one on day 4; therefore, subsequent CASUS measurements were unavailable. The linear mixed-effects model showed a group-by-day interaction [$F(4, 235)=5.28$; $p<0.001$], indicating different average score patterns for survivors and non-survivors. In the exploratory unadjusted Welch's t-tests, scores were 5.6±3.6 versus 8.7±5.5 on day 4 (p=0.011) and 4.6±3.2 versus 9.4±6.5 on day 5 (p<0.001), respectively (Table 3, Figure 1).

The effective sample sizes for ROC analyses on postoperative days 1-5 were 63, 62, 62, 60, and 59, respectively. AUCs were 0.584 on day 1 (p=0.281), 0.575 on day 2 (p=0.337), 0.620 on day 3 (p=0.126), 0.664 on day 4 (p=0.037), and 0.729 on day 5 (p=0.004). The exploratory Youden cut-off for day 4 was >8.5 (sensitivity 53.8%, specificity 76.2%), and the day 5 cut-off was >8.5 (sensitivity 55.3%, specificity 90.5%). Day 5 CASUS, therefore, showed moderate discrimination but limited sensitivity in this selected cohort. EuroSCORE II did not show evidence of discrimination (AUC =0.413; p=0.265). All day-specific analyses and cut-offs were internally derived and should be considered exploratory (Table 4, Figure 2).

The day-5 complete-case analyses included 59 patients (21 survivors and 38 non-survivors). In univariable analysis, day 5 CASUS was associated with mortality (OR 1.204, 95% CI 1.055-1.375; p=0.006), whereas age (OR 1.004, 95% CI 0.941-1.070; p=0.913), CPB duration (OR 1.003, 95% CI 0.994-1.013; p=0.490), and EuroSCORE II (OR 0.836, 95% CI 0.621-1.126; p=0.239) were not associated with mortality. After adjustment, day 5 CASUS remained associated with higher odds of in-hospital mortality (adjusted OR 1.245, 95% CI 1.076-1.439; p=0.003); each one-point increase corresponded to a 24.5% increase in the adjusted odds, not the probability, of mortality. The Hosmer-Lemeshow test did not indicate a gross lack of fit (chi-square =8.382, df=8; p=0.397), but calibration cannot be established from this test in a small sample. Nagelkerke R^2 was 0.309, and the model was neither internally nor externally validated (Table 5, Figure 3).

Table 1. Demographic, preoperative, and procedural characteristics [mean ± SD, n (%)]			
Variable	**Survivors (n=21)**	**Non-survivors (n=42)**	**p**
Age (years)	66.1±5.9	66.1±9.7	0.980
Sex (M/F)	10/11 (47.6/52.4)	27/15 (64.3/35.7)	0.270
BMI (kg/m^2)	28.1±5.4	27.4±4.4	0.590
Diabetes mellitus	12 (57.1)	28 (66.7)	0.580
Hypertension	15 (71.4)	28 (66.7)	0.780
CAD	19 (90.5)	34 (81.0)	0.470
COPD	4 (19.0)	5 (11.9)	0.470
CKD	5 (23.8)	7 (16.7)	0.510
EF (%)	46.0±10.4	47.3±10.7	0.640
EuroSCORE II	6.1±2.5	5.5±1.3	0.220
Surgical procedure type, n (%)			**0.403**
CABG	10 (47.6)	17 (40.5)	
Combined surgery	3 (14.3)	14 (33.3)	
Valve surgery	5 (23.8)	6 (14.3)	
Aortic surgery	3 (14.3)	5 (11.9)	

Student's t-test/Mann-Whitney U test, chi-square test, or exact test, as appropriate. The overall surgical procedure distribution was compared using the Fisher-Freeman-Halton exact test. BMI: Body mass index, CABG: Coronary artery bypass grafting, CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, CKD: Chronic kidney disease, EF: Ejection fraction, SD: Standard deviation, M/F: Male/Female

Table 2. Intraoperative characteristics (mean ± SD)			
Variable	**Survivors (n=21)**	**Non-survivors (n=42)**	**p**
Cross-clamp time (min)	82.0±46.5	87.5±47.6	0.660
CPB time (min)	130.2±43.7	139.8±64.6	0.540
Total surgical time (min)	205.7±41.4	216.0±40.7	0.353

Student's t-test. CC: Cross-clamp, CPB: Cardiopulmonary bypass, Min: Minutes, SD: Standard deviation

Table 3. Daily postoperative CASUS scores according to in-hospital mortality status (mean ± SD)			
Postoperative day	**Survivors**	**Non-survivors**	**p†**
Day 1 (n=21/42)	9.3±3.2	10.3±3.6	0.243
Day 2 (n=21/41)	7.0±3.7	8.5±5.0	0.209
Day 3 (n=21/41)	6.2±4.4	8.8±6.1	0.061
Day 4 (n=21/39)	5.6±3.6	8.7±5.5	0.011*
Day 5 (n=21/38)	4.6±3.2	9.4±6.5	<0.001**
Group × day interaction, p‡			<0.001**

†Welch's t-test. Day-specific p values are unadjusted and exploratory. ‡Group-by-day interaction from the linear mixed-effects model: $F(4, 235)=5.28$; $p<0.001$. Later observations unavailable because of death were not imputed. Sample sizes are presented as survivors/non-survivors. *$p<0.05$; **$p<0.001$. SD: Standard deviation

DISCUSSION

Among the 63 patients in this single-center PMV cohort, higher CASUS values later in the postoperative course were associated with mortality before discharge. Exploratory between-group differences emerged on postoperative days 4 and 5; in the adjusted day-5 model, the association with CASUS persisted after accounting for age, CPB duration, and EuroSCORE II. However, the day-5 AUC of 0.729 represents moderate, rather than high, discrimination, and the sensitivity at the exploratory >8.5 cut-off was 55.3%. The finding supports an association with evolving postoperative severity in this restricted cohort, and does not establish an early prediction tool or a definitive clinical threshold.

Earlier work has generally assessed CASUS in larger cardiac surgical cohorts and reported greater discrimination than we observed. In studies of acute aortic dissection and mixed postoperative patient populations, CASUS supported the

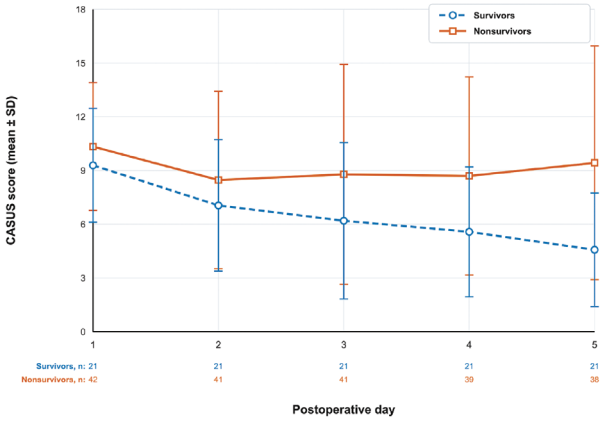


Figure 1. Daily postoperative CASUS scores according to in-hospital mortality status. Points represent observed group means and error bars represent standard deviations. Day-specific sample sizes (survivors/non-survivors) for postoperative days 1-5 were 21/42, 21/41, 21/41, 21/39, and 21/38, respectively. The reduction in available observations among non-survivors reflects deaths that occurred before later measurements. Lines connect group-level means for visualization only, and do not represent patient-specific trajectories. Daily between-group comparisons, using Welch's t-test, were exploratory and not adjusted for multiple testing.

CASUS: Cardiac surgery score, SD: Standard deviation

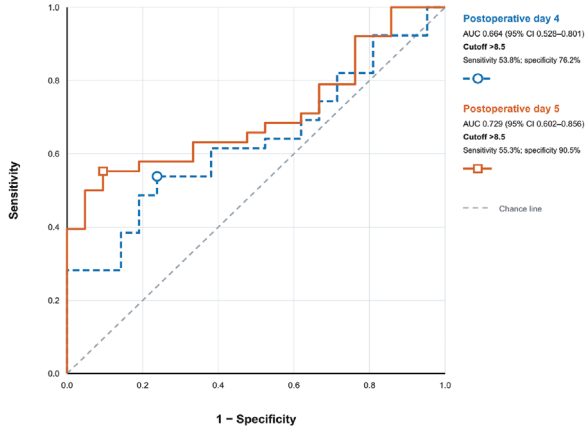


Figure 2. Receiver operating characteristic curves for postoperative day 4 and day 5 CASUS scores for discriminating in-hospital mortality. The day-4 analysis included 60 patients (21 survivors and 39 non-survivors): AUC =0.664 (95% CI 0.528-0.801), p=0.037; the Youden cut-off was >8.5, with 53.8% sensitivity and 76.2% specificity. The day-5 analysis included 59 patients (21 survivors and 38 non-survivors). AUC =0.729 (95% CI 0.602-0.856), p=0.004; the Youden cut-off was >8.5, with 55.3% sensitivity and 90.5% specificity. The dashed diagonal denotes chance discrimination. These internally derived cut-offs are exploratory and center-specific, rather than validated clinical thresholds.

AUC: Area under the curve, CASUS: Cardiac surgery score, CI: Confidence interval

assessment of mortality and sometimes outperformed general ICU scores.[9,10] A study confined to CABG patients in Türkiye found higher CASUS values among patients who were reintubated or who died within 30 days; for 7-day mortality and prolonged ICU stay, its AUCs exceeded those of APACHE II.[11] On postoperative day 2, logistic CASUS provided moderate discrimination for predicting prolonged ICU stay.[12] A separate postoperative cohort also reported that CASUS discriminated mortality risk after cardiac surgery.[13] Our estimates were less favorable (day 5 AUC: 0.729; sensitivity: 55.3%). The proximity of our exploratory cut-off to values reported in other cohorts is descriptive only and neither validates the cut-off nor supports stand-alone clinical use.

Several factors may explain the lower discrimination compared with prior validation studies: restriction to patients with PMV; a narrowed, more severe risk spectrum; small sample size; high outcome prevalence; use of day-specific, rather than maximum or summary, scores; different mortality definitions; and loss of later measurements due to death. The absence of uncomplicated patients is especially important, because the performance in a restricted PMV cohort cannot be extrapolated to the full postoperative cardiac surgery population.

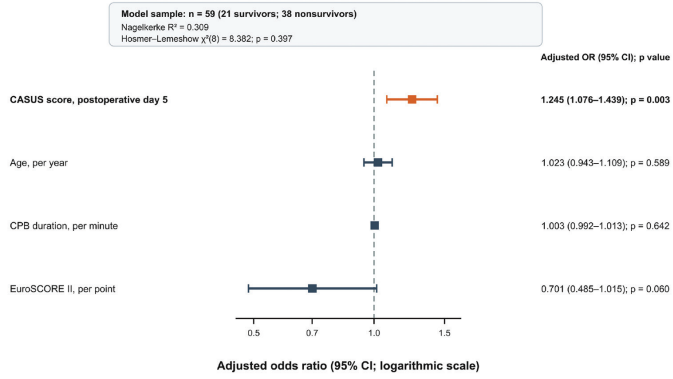


Figure 3. Adjusted associations with in-hospital mortality among patients with an available postoperative day 5 CASUS score. Odds ratios and 95% confidence intervals were estimated from a multivariable binary logistic regression model that included postoperative day 5 CASUS, age, cardiopulmonary bypass duration, and EuroSCORE II. The model included 59 patients (21 survivors and 38 non-survivors). Four non-survivors who died before postoperative day 5 had no day 5 score and were not included. Points represent adjusted odds ratios, horizontal lines represent 95% confidence intervals, and the dashed vertical line denotes OR =1. The CASUS estimate is highlighted in orange. The Hosmer-Lemeshow test did not indicate gross lack of fit; however, calibration remains uncertain because of the small sample size and the absence of validation.

CASUS: Cardiac surgery score, CI: Confidence interval, CPB: Cardiopulmonary bypass, OR: odds ratio

Table 4. Exploratory ROC analyses of CASUS and EuroSCORE II for in-hospital mortality					
Variable (n)	**AUC (95% CI)**	**Cut-off**	**Sensitivity**	**Specificity**	**p†**
CASUS, day 1 (63)	0.584 (0.432-0.735)	-	-	-	0.281
CASUS, day 2 (62)	0.575 (0.429-0.720)	-	-	-	0.337
CASUS, day 3 (62)	0.620 (0.478-0.761)	-	-	-	0.126
CASUS, day 4 (60)	0.664 (0.528-0.801)	>8.5	53.8%	76.2%	0.037*
CASUS, day 5 (59)	0.729 (0.602-0.856)	>8.5	55.3%	90.5%	0.004*
EuroSCORE II (63)	0.413 (0.248-0.578)	-	-	-	0.265

†Versus H0: AUC =0.5. Day-specific analyses used available cases. No adjustment for multiple testing was applied. Youden cut-offs were derived and evaluated in the same dataset and are exploratory, center-specific estimates. *p<0.05. AUC: Area under the curve, CI: Confidence interval, ROC: Receiver operating characteristic, CASUS: Cardiac surgery score

Table 5. Univariable and multivariable logistic regression analyses for in-hospital mortality (complete-case n=59; 21 survivors and 38 non-survivors)					
Variable	**n**	**Unadjusted OR (95% CI)**	**p**	**Adjusted OR (95% CI)**	**p**
CASUS, day 5	59	1.204 (1.055-1.375)	0.006	1.245 (1.076-1.439)	0.003
Age	59	1.004 (0.941-1.070)	0.913	1.023 (0.943-1.109)	0.589
CPB duration	59	1.003 (0.994-1.013)	0.490	1.003 (0.992-1.013)	0.642
EuroSCORE II	59	0.836 (0.621-1.126)	0.239	0.701 (0.485-1.015)	0.060

The day-5 CASUS was the predictor of interest. Age, CPB duration, and EuroSCORE II were included on clinical grounds as adjustment covariates, rather than being selected on the basis of univariable significance. Nagelkerke R^2=0.309; Hosmer-Lemeshow χ^2=8.382, df=8, p=0.397. A non-significant Hosmer-Lemeshow test does not establish good calibration in this small sample. OR: Odds ratio, CI: Confidence interval, CPB: Cardiopulmonary bypass, CASUS: Cardiac surgery score

External validation studies indicate that CASUS performance and calibration can vary across health-care systems and may require local updating.[8,14] This supports independent validation rather than the adoption of a center-specific value derived from and tested in the same cohort. Our day 4 and day 5 cut-offs may be optimistic and should be regarded only as exploratory estimates.

The group-by-day interaction indicates that average postoperative CASUS patterns differed between survivors and non-survivors. It does not establish the trajectory itself as an independent predictor because the regression model included day-5 CASUS rather than an individual change score or slope. EuroSCORE II also had a confidence interval for the AUC that included 0.5 in this selected cohort. The inverse EuroSCORE II estimate was imprecise and should be regarded as a potentially unstable finding in this small, selected cohort rather than as evidence of a protective association. Accordingly, this cohort-specific estimate cannot be used to judge EuroSCORE II performance in broader populations. Combined preoperative and postoperative approaches remain of interest.[15]

Prior research that directly analyzed delta CASUS found that changes over time can be prognostically informative.[16] Our analysis did not model delta CASUS or a patient-specific slope. Such direct trajectory measures should be examined prospectively in a larger cohort before the postoperative CASUS course can be considered an independently validated predictor.

All patients in this study were selected because they developed PMV; the PaO_2/FiO_2 ratio is a component of CASUS. Respiratory dysfunction, therefore, contributed to both cohort selection and the measured score. This conceptual overlap may affect the score distribution and apparent performance, and may further limit extrapolation to patients without PMV. Findings in other critically ill subgroups, including patients receiving extracorporeal life support, provide clinical context but do not remove this limitation.[17]

Dynamic risk assessment may help characterize changes in postoperative severity.[18] Nevertheless, a CASUS value obtained on day 5 is relatively late and may reflect persistent organ dysfunction, complications, and treatment response among patients who survived long enough to be assessed. It should not be interpreted as an early prognostic marker, and clinical use requires prospective validation.

Study Limitations

Several limitations should be considered. Because the data were derived from a retrospective, single-center PMV cohort, generalizability to uncomplicated postoperative patients is limited. Fifty-five of 118 PMV patients (46.6%) were excluded because the data required for the postoperative day 1 CASUS score were incomplete, potentially introducing additional selection bias. All eligible patients from the prespecified interval were analyzed; therefore, no pre-specified sample-size target was set. The day-5 regression included only 59 patients (21 survivors and 38 non-survivors) and used four predictors; estimates may be unstable, and no internal or external validation was performed. Four non-survivors had no later scores because they died before day 5.

Informative dropout and conditioning on availability at day 5 may have affected longitudinal and regression estimates; a mixed-effects model does not eliminate this bias. Because the study did not model an individual delta CASUS or slope, the trajectory was not established as an independent predictor. No multiplicity adjustment was applied; day-specific comparisons and ROC analyses were exploratory, and internally derived cut-offs may be optimistic and center-specific. Selection on PMV also overlaps conceptually with the PaO_2/FiO_2 component of CASUS. Finally, the comparison was limited to EuroSCORE II and did not include postoperative comparators such as SOFA or APACHE II.

CONCLUSION

In this selected single-center cohort of open-heart surgery patients with PMV lasting longer than 24 hours, a higher postoperative day 5 CASUS score was associated with increased odds of in-hospital mortality and demonstrated moderate discriminatory ability. The day-5 score should be interpreted as a marker of evolving postoperative severity in patients for whom the measurement was available, rather than as an early prognostic marker. Differences in average daily score patterns did not indicate that the trajectory was an independent predictor. The exploratory, center-specific cut-offs require validation in larger multicenter cohorts before clinical use.

Ethics

Ethics Committee Approval: The study was approved by the Medical Sciences Ethics Committee of the University of Health Sciences Bursa Yüksek İhtisas Training and Research Hospital on March 11, 2026 (decision no: 2024-TBEK 2026/03-02).

Informed Consent: Owing to the retrospective design, the requirement for informed consent was waived by the ethics committee.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.O., O.S.A., N.B.S., Concept: A.O., M.Ö., Design: O.S.A., M.Ö., Data Collection or Processing: A.B.T., C.S.A., Analysis or Interpretation: N.B.S., T.O., Literature Search: A.O., Ü.K., T.O., Writing: A.O., C.S.A., A.K.A., Ü.K., T.O.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Squiccimarro E, Stasi A, Lorusso R, Paparella D. Narrative review of the systemic inflammatory reaction to cardiac surgery and cardiopulmonary bypass. Artif Organs. 2022;46:568-77.
2. Wang Q, Tao Y, Zhang X, et al. The incidence, risk factors, and hospital mortality of prolonged mechanical ventilation among cardiac surgery patients: a systematic review and meta-analysis. Rev Cardiovasc Med. 2024;25:409.
3. Aksoy R, Karakoc AZ, Cevirme D, et al. Predictive factors of prolonged ventilation following cardiac surgery with cardiopulmonary bypass. Braz J Cardiovasc Surg. 2021;36:780-7.
4. Liu Q, Chen P, Wang W, et al. A novel scoring model for predicting prolonged mechanical ventilation in cardiac surgery patients: development and validation. Front Cardiovasc Med. 2025;12:1573874.
5. Mastroiacovo G, Bonomi A, Ludergnani M, et al. Is EuroSCORE II still a reliable predictor for cardiac surgery mortality in 2022? A retrospective study study. Eur J Cardiothorac Surg. 2022;64:ezad294.
6. Howell NJ, Head SJ, Freemantle N, et al. The new EuroSCORE II does not improve prediction of mortality in high-risk patients undergoing cardiac surgery: a collaborative analysis of two European centres. Eur J Cardiothorac Surg. 2013;44:1006-11; discussion 1011.
7. Doerr F, Badreldin AM, Heldwein MB, et al. A comparative study of four intensive care outcome prediction models in cardiac surgery patients. J Cardiothorac Surg. 2011;6:21.
8. Wilson B, Tran DTT, Dupuis JY, McDonald B. External validation and updating of the cardiac surgery score for prediction of mortality in a cardiac surgery intensive care unit. J Cardiothorac Vasc Anesth. 2019;33:3028-34.
9. Gürcü ME, Külahcioglu S, Karaca Baysal P, et al. Comparison of sequential organ failure assessment score and cardiac surgery score systems for mortality prediction after emergency acute aortic dissection surgery. Ulus Travma Acil Cerrahi Derg. 2022;28:1298-304.
10. Badreldin AM, Doerr F, Ismail MM, et al. Comparison between sequential organ failure assessment score (SOFA) and cardiac surgery score (CASUS) for mortality prediction after cardiac surgery. Thorac Cardiovasc Surg. 2012;60:35-42.
11. Uğur S, Acarel M, Yapıcı N. CASUS and APACHE II score in predicting mortality after coronary artery bypass grafting. Turk Gogus Kalp Damar Cerrahisi Derg. 2023;31:343-51.
12. Pimentel MF, Soares MJF, Murad JA Junior, et al. Predictive factors of long-term stay in the ICU after cardiac surgery: logistic CASUS score, serum bilirubin dosage and extracorporeal circulation time. Braz J Cardiovasc Surg. 2017;32:367-71.
13. Raut S, Hussain A, Ariyaratnam P, et al. Validation of cardiac surgery score (CASUS) in postoperative cardiac patients. Semin Cardiothorac Vasc Anesth. 2020;24:304-12.
14. Singh A, Liang C, Mick SL, Udeh C. External validation of the cardiac surgery score in a quaternary hospital in the United States of America. J Intensive Care Med. 2022;37:1318-27.
15. Doerr F, Heldwein MB, Bayer O, et al. Combination of European System for Cardiac Operative Risk Evaluation (EuroSCORE) and cardiac surgery score (CASUS) to improve outcome prediction in cardiac surgery. Med Sci Monit Basic Res. 2015;21:172-8.
16. Badreldin AM, Kroener A, Heldwein MB, et al. Prognostic value of daily cardiac surgery score (CASUS) and its derivatives in cardiac surgery patients. Thorac Cardiovasc Surg. 2010;58:392-7.
17. Hofmann B, Gmelin MJ, Metz D, et al. Cardiac surgery score (CASUS) improves outcome prediction in patients treated with extracorporal life support (ECLS). Perfusion. 2018;33:36-43.
18. Diaz-Soto JC, Couture EJ, Nabzdyk CGS. Quo vadis cardiac surgery risk score? From predicting to impacting outcomes in cardiac surgery. J Cardiothorac Vasc Anesth. 2022;36:995-7.

DOI: 10.4274/forbes.galenos.2026.10327
Forbes J Med 2026;7:122-7

Original Article

Radiological Validation and Explainability of a Deep Learning Model for Vertebral Compression Fracture Detection in Geriatric Spine Imaging

Anıl ÖZGÜR[1], Ali Murat BAŞAK[2]

[1]İnegöl State Hospital, Clinic of Orthopedics and Traumatology, Bursa, Türkiye
[2]University of Health Sciences Türkiye, Gülhane Training and Research Hospital, Clinic of Orthopedics and Traumatology, Ankara, Türkiye

Cite as: Özgür A, Başak AM. Radiological validation and explainability of a deep learning model for vertebral compression fracture detection in geriatric spine imaging. Forbes J Med 2026;7:122-122

ABSTRACT

Objective: Osteoporotic vertebral compression fractures are common in older adults and can significantly impair quality of life. Early and accurate detection of these fractures remains clinically challenging. This study aimed to develop and evaluate an explainable deep learning framework for the automated detection and classification of vertebral compression fractures on lumbar computed tomography (CT) images.

Methods: Eighty anonymized lumbar CT scans from the VerSe2020 and SpineWeb open-access datasets were analyzed. Vertebrae were segmented, and morphological features, including height loss and posterior cortical integrity, were extracted. A dual-stage deep learning pipeline combining YOLOv8 for fracture localization and EfficientNet-B0 for classification was developed to predict both the presence of fractures and their surgical eligibility. Model performance was evaluated using the following metrics: accuracy, F1 score, sensitivity, specificity, and receiver operating characteristic-area under the curve (ROC-AUC).

Results: The model achieved an accuracy of 68%, an F1 score of 63%, a sensitivity of 60%, a specificity of 65%, and an ROC-AUC of 0.66. Additionally, the model's decision-making focus on morphologically relevant zones suggests its potential use for automated screening and prioritization in radiological workflows.

Conclusion: This artificial intelligence-based framework may support early diagnosis and preoperative assessment of vertebral fractures in geriatric patients and has the potential to assist clinical decision-making.

Keywords: Vertebral fracture, artificial intelligence, radiological imaging, deep learning

Received: 17.07.2026
Accepted: 15.09.2026
Publication Date: 02.10.2026

Corresponding Author:
Anıl ÖZGÜR
İnegöl State Hospital, Clinic of Orthopedics and Traumatology, Bursa, Türkiye
dr.anilozgur@gmail.com
ORCID ID: 0000-0002-5877-0628

INTRODUCTION

Osteoporotic vertebral compression fractures are common in elderly and postmenopausal patients and are associated with persistent pain, spinal deformity, and reduced quality of life.[1] Early radiological detection is particularly important in elderly patients, as delayed diagnosis may contribute to progressive vertebral collapse and functional deterioration. Percutaneous kyphoplasty is widely used to restore vertebral height and relieve pain in selected patients; however, its indications should be considered carefully because of variable outcomes and procedure-related complications, including new fractures, cement leakage, and spinal canal compression.[2,3]

Artificial intelligence (AI)-based approaches in orthopedic imaging have rapidly expanded in recent years, enabling automated fracture detection and classification using computed tomography (CT) and magnetic resonance imaging data.[4,5] Beyond imaging analysis, large language model-based systems such as ChatGPT have also demonstrated increasing potential in orthopedic decision-

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

support processes and patient-oriented clinical guidance.[6] In addition, digital health research has shown that online orthopedic information-seeking behavior demonstrates significant temporal and seasonal variation, highlighting the growing interaction between AI, digital epidemiology, and orthopedic practice.[7] Deep learning models, including three-dimensional (3D) convolutional neural networks and Mask R-CNN, have shown high accuracy in vertebral segmentation and radiological interpretation.[8] In addition, AI models have been applied to predict clinical outcomes such as fracture risk, cement leakage, and postoperative complications, supporting more personalized treatment strategies.[9-11]

Despite these advances, few studies have combined automated vertebral fracture detection with kyphoplasty-related clinical assessment using AI.[12] Therefore, this study aimed to detect vertebral compression fractures in open-access CT datasets and to evaluate their clinical relevance using a deep learning-based approach.

METHODS

Study Design and Ethical Approval Status

This study is a non-interventional, AI-based analysis that uses anonymized, open-access CT datasets (VerSe 2020 and SpineWeb) to detect vertebral compression fractures and evaluate their clinical relevance. The primary objective was to develop a model capable of identifying fractures and supporting preoperative assessment. No direct patient data, clinical records, or identifiable information were used in the study. All imaging data were obtained from publicly available sources and processed solely for research purposes. The study was conducted in accordance with the Declaration of Helsinki and Good Clinical Practice guidelines. As the analysis was performed on fully de-identified datasets without involvement of human subjects, approval from the ethics committee was not required. The study design reflects a technical and observational approach aimed at supporting clinical decision-making.

Image Sources and Data Set Properties

In this study, open-access CT images from the VerSe 2020 and SpineWeb datasets were used to identify vertebral fractures and assess their clinical relevance. The VerSe dataset includes vertebral segmentation data from thoracolumbar vertebrae, while SpineWeb provides CT images of compression fractures across different age groups. A total of 80 patients (approximately 150-200 sagittal slices per case) were analyzed. Cases were included if sagittal CT images provided adequate visualization of the vertebral bodies and were accompanied by reference annotations that were suitable for model training and evaluation. Cases with incomplete vertebral visualization, severe image artifacts, or insufficient annotation data were excluded. The 80 subjects, therefore, represented eligible cases meeting these imaging-based criteria rather than a random sample. Images were obtained in DICOM format and preprocessed, and ground-truth segmentation masks were derived from reference annotations. Morphological parameters, including vertebral body height, fracture severity (Genant classification), and posterior cortical integrity, were extracted for analysis.

Artificial Intelligence Model and Training Process

The proposed model was a dual-stage deep-learning framework for vertebral compression fracture detection and radiological validation. First, a YOLOv8-based module was used to localize and segment vertebral regions from sagittal CT images; standard preprocessing and data augmentation were applied to improve generalizability. In the second stage, an EfficientNet-B0 classifier classified the segmented vertebrae as fractured or intact. The model captured key morphological features such as vertebral height loss, cortical irregularity, and posterior wall disruption.

Model training was performed using the PyTorch framework. A total of 5,600 sagittal CT images, obtained from 80 subjects, were included in the image-level classification analysis and constituted the effective sample for model development and evaluation. The images were divided into training, validation, and test sets at a ratio of 70%, 15%, and 15%, respectively. Individual vertebrae were not treated as independent subject-level observations; therefore, the effective sample size was defined by the number of images included in the classification analysis rather than by a separate vertebra-level count. Model performance was evaluated using accuracy, sensitivity, specificity, F1-score, and ROC-AUC metrics. To enhance interpretability, Grad-CAM visualizations were generated and qualitatively assessed by two independent musculoskeletal radiologists for anatomical consistency.

Radiological Criteria for Fracture Severity

Fracture severity was defined using commonly accepted radiological criteria, including >40% anterior vertebral height loss, posterior wall disruption, segmental kyphotic angle >10°, and loss of posterior cortical continuity. These criteria were used to evaluate the model's ability to identify morphologically severe fractures rather than guide treatment decisions. To assess anatomical validity and inter-observer consistency, two independent musculoskeletal radiologists reviewed a 20% subset of the cases as a quality-control sample. This subset assessment was performed to verify the anatomical consistency of

the predefined radiological criteria, rather than to provide a second radiological evaluation of the entire dataset. Substantial inter-observer agreement (κ=0.82) was observed, supporting the consistency of the radiological assessment within the reviewed subset.

Evaluation and Performance Analysis

Model performance was evaluated on the test dataset using standard classification metrics, including accuracy, sensitivity, specificity, F1-score, and receiver operating characteristic-area under the curve (ROC-AUC). The ROC curve was used to assess the model's ability to discriminate between fractured and non-fractured vertebrae, and an optimal decision threshold was determined. In addition, fracture identification performance was evaluated using the true positive rate, false negative rate, and false positive rate. Segmentation performance was also evaluated using the mean Intersection over Union (mIoU).

RESULTS

The developed deep learning framework was successfully applied to the analysis of vertebral compression fractures using open-access CT datasets. A total of 5,600 sagittal slices obtained from 80 anonymized subjects were processed using the dual-stage model architecture. In the first stage, the YOLOv8-based localization network effectively segmented vertebral boundaries across the entire dataset, reaching a mIoU score of 0.91±0.04, which confirmed accurate regional detection. Regional detection accuracy was verified by comparing the predicted vertebral regions with reference vertebral segmentation annotations from open-access datasets, which served as the ground truth for IoU calculation. Only a small proportion of slices (1.6%) were excluded due to incomplete vertebral visualization or motion-related artifacts.

Following segmentation, the EfficientNet-B0 classifier was trained to differentiate normal from fractured vertebrae, based on predefined morphological criteria. The model achieved an accuracy of 68% and an F1 score of 0.63, indicating balanced classification performance. Sensitivity and specificity were 60% and 65%, respectively, demonstrating a moderate ability to identify fractured and intact vertebrae. ROC analysis yielded an AUC of 0.66, reflecting moderate discriminatory performance (Table 1).

Based on radiological severity criteria, the model correctly classified 49% of vertebrae with severe morphological features. False-negative and false-positive rates were 22% and 17%, respectively; these were mainly due to subtle fractures and degenerative changes that mimicked compression. Overall agreement between AI predictions and expert radiological assessments was 52%, indicating moderate consistency (Table 2).

To further understand model performance, an error distribution analysis was conducted. The most frequent limitation was the omission of marginal (mild) fractures, accounting for 28% of total errors. Ambiguity in vertebral border segmentation constituted 19%, while undetected posterior wall damage accounted for 14%. Image quality variation among the open datasets accounted for 21% of the observed performance inconsistencies, while model overfitting accounted for 11%.

The ROC curve illustrates the model's discriminative ability to distinguish fractured from non-fractured vertebrae (Figure 1). The curve's deviation from the ideal diagonal indicates that the model has limited but significant discriminative ability, suggesting that performance could be further improved with larger and more diverse datasets. The optimal threshold for binary classification was determined to be 0.53, balancing sensitivity and specificity in future applications.

DISCUSSION

In the present study, the proposed deep learning framework demonstrated moderate diagnostic performance in detecting vertebral compression fractures on open-access CT images. Given the clinical importance of timely fracture recognition, these findings suggest a potential supportive role for AI-assisted radiological assessment and emphasize the limitations of the current model. Importantly, the framework was designed to identify radiologically relevant fracture features rather than to determine treatment indications, which remain dependent on comprehensive clinical and radiological evaluation.[1,12]

AI-based approaches for vertebral fracture detection have gained increasing attention in recent years. Deep learning models trained on CT data have demonstrated high diagnostic performance for fracture localization and classification under controlled conditions.[13,14] Some studies have reported ROC-AUC values exceeding 0.90 using 3D convolutional architectures and institution-specific datasets, indicating that automated systems can achieve expert-level accuracy in optimal settings.[14,15] In comparison, the dual-stage deep learning framework used in the present study, combining YOLOv8 for vertebral localization and EfficientNet-B0 for classification, achieved an ROC-AUC of 0.66, indicating moderate discriminative performance and lower diagnostic accuracy than the >0.90 values reported under more controlled conditions. This finding is more consistent with the moderate performance reported for models based on heterogeneous open-access datasets. In contrast to these high-performing models, approaches relying exclusively on open-access datasets generally report more moderate results due to heterogeneity in image quality, acquisition protocols, and annotation

Table 1. Advanced model performance metrics

Metric	Definition	Formula	Model result (%)	Interpretation
Accuracy	Model accuracy shows the proportion of samples correctly classified among all predictions	(TP + TN)/(TP + TN + FP + FN)	68	Overall accuracy is at an average level
F1 score	Balance between precision and sensitivity	2× (Precision × sensitivity)/(precision + sensitivity)	63	Balance between sensitivity and scope achieved
Sensitivity	The proportion of true positives	TP/(TP + FN)	60	The model captures positives at a rate of 60%
Specificity	The proportion of true negatives	TN/(TN + FP)	65	Moderate ability to differentiate negatives
ROC-AUC	Ability to differentiate true positive/negative	AUC	66	The model's classification capacity is limited but significant

TP: True positive, TN: True negative, FP: False positive, FN: False negative, ROC-AUC: Receiver operating characteristic-area under the curve

Table 2. Model performance in identifying radiologically severe vertebral compression fractures

Criterion	Definition	Percentage (%)	Clinical significance	Interpretation
True positive	Radiologically severe fractures correctly identified by the model	49	It has the potential to contribute to treatment	The model correctly identified nearly half of kyphoplasty candidates
False negative	Radiologically severe fractures not identified by the model	22	They may miss the opportunity for treatment.	Lack of sensitivity is a serious limitation.
False positive	Cases incorrectly classified by the model as radiologically severe fractures	17	Increases the risk of unnecessary operations	An important source of error affecting the specificity
Overall agreement	Agreement between model predictions and expert radiological assessment	52	Indicates overall reliability	The decision-making consistency of the model is moderate
Overclassification rate	Cases incorrectly classified as radiologically severe by the model	17	Clinical safety should be questioned	A cautious approach should be taken to clinical decision support

standards.[4,5] Studies that combine image-based analysis with quantitative morphological features, such as vertebral height loss ratios, have shown improved sensitivity and specificity compared with image-only models.[15,16] These findings suggest that incorporating structured radiological parameters can substantially enhance fracture detection performance. Importantly, the model presented in this study focused on anatomically relevant regions, specifically cortical disruption and vertebral height loss, thereby further supporting its radiological validity.

Beyond fracture detection, machine learning methods have been increasingly applied to predict clinical outcomes related to vertebral augmentation procedures, including postoperative pain, cement leakage, and fracture recurrence.[11,17,18] While these predictive models typically incorporate demographic and clinical variables in addition to imaging data, they highlight the broader potential of AI to support orthopedic decision-making across multiple stages of patient management.

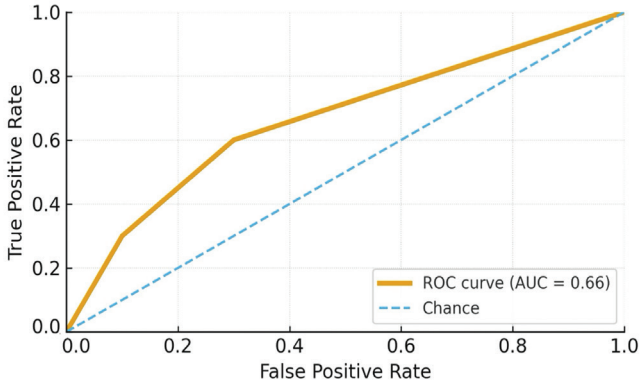


Figure 1. Receiver operating characteristic curve illustrating the model's ability to discriminate between fractured and non-fractured vertebrae (AUC =0.66)

ROC: Receiver operating characteristic, AUC: Area under the curve

Model interpretability is essential for clinical adoption of AI systems. Visualization techniques, such as Grad-CAM, improve transparency by highlighting regions that influence model decisions.[19,20] In this study, Grad-CAM analysis demonstrated that the model focused on morphologically relevant vertebral regions, supporting the reliability of its predictions.

Despite its potential, the present study has several limitations. The model was trained using a limited number of open-access imaging datasets, and heterogeneity in image quality, acquisition protocols, and annotation standards may restrict its generalizability across different patient populations. In addition, the absence of demographic and clinical variables, such as age, sex, and osteoporosis severity, limits the clinical applicability of the current framework. The error analysis also demonstrated reduced performance in detecting subtle fractures and posterior cortical disruption and revealed segmentation-related uncertainties. These findings indicate that the current model is not sufficient to replace expert radiological interpretation. Future studies should therefore focus on larger and more diverse datasets, integration of relevant clinical parameters, and refinement of the model architecture, particularly for assessing posterior wall involvement.

The proposed framework may be used as a pre-screening tool within radiological workflows, helping to identify suspicious vertebral regions and to reduce diagnostic delays in geriatric patients.[5,21] Such systems should be considered supportive tools rather than replacements for expert interpretation.[22] Overall, these findings demonstrate the feasibility of explainable AI models for vertebral fracture assessment and their potential role in orthopedic imaging.

CONCLUSION

The proposed dual-stage deep learning framework demonstrated potential as a supportive tool for radiological detection of vertebral compression fractures in CT images. Its ability to identify morphologically relevant changes in vertebrae may facilitate early detection and prioritization within radiological workflows. Further prospective, multicenter validation is required before clinical implementation.

Ethics

Ethics Committee Approval: As the analysis was performed on fully de-identified datasets without involvement of human subjects, approval from the ethics committee was not required.

Informed Consent: No direct patient data, clinical records, or personally identifiable information were used in this study. All imaging data were obtained from publicly available sources and processed solely for research purposes. Therefore, informed consent was not required.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.Ö., A.M.B., Concept: A.Ö., A.M.B., Design: A.Ö., A.M.B., Data Collection or Processing: A.Ö., A.M.B., Analysis or Interpretation: A.Ö., A.M.B., Literature Search: A.Ö., A.M.B., Writing: A.Ö., A.M.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Ma Y, Lu Q, Yuan F, Chen H. Comparison of the effectiveness of different machine learning algorithms in predicting new fractures after PKP for osteoporotic vertebral compression fractures. J Orthop Surg Res. 2023;18:62.
2. Chettrit D, Meir T, Lebel H, et al. 3D convolutional sequence to sequence model for vertebral compression fractures İdentification in CT. 2020.
3. Nicolaes J, Raeymaeckers S, Robben D, et al. Detection of vertebral fractures in CT using 3D convolutional neural networks. 2020.
4. Paik S, Park J, Hong JY, Han SW. Deep learning application of vertebral compression fracture detection using mask R-CNN. Sci Rep. 2024;14:16308.
5. Kim YR, Yoon YS, Cha JG. Opportunistic screening for acute vertebral fractures on a routine abdominal or chest computed tomography scans using an automated deep learning model. Diagnostics. 2024;14:781.
6. Karadamar ÖL, Aydilek A. Evaluation of surgical recommendation algorithms in hallux valgus cases using ChatGPT models: an artificial intelligence approach based on 50 simulated scenarios. Acta Orthop Traumatol Turc. 2026;60:e25580.

7. Karadamar ÖL, Kara M, Özdin T, Aydilek A, Özgür A. Search trends of orthopedic terms in Turkey: a five-year time series analysis based on Google Trends. Journal of Multidisciplinary Orthopaedic Surgery. 2025;1:22-6.
8. Lee J, Kim M, Park H, et al. Enhanced detection performance of acute vertebral compression fractures using a hybrid deep learning and traditional quantitative measurement approach: beyond the limitations of genant classification. Bioengineering (Basel). 2025;12:64.
9. Namireddy SR, Gill SS, Peerbhai A, et al. Artificial intelligence in risk prediction and diagnosis of vertebral fractures. Sci Rep. 2024;14:30560.
10. Seo JW, Lim SH, Jeong JG, Kim YJ, Kim KG, Jeon JY. A deep learning algorithm for automated measurement of vertebral body compression from X-ray images. Sci Rep. 2021;11:13732.
11. Hu YL, Wang PY, Xie ZY, et al. Interpretable machine learning model to predict bone cement leakage in percutaneous vertebral augmentation for osteoporotic vertebral compression fracture based on SHapley additive exPlanations. Global Spine J. 2025;15:689-701.
12. Zhao Y, Bo L, Chen X, et al. Evaluation and analysis of risk factors for adverse events of the fractured vertebra post-percutaneous kyphoplasty: a retrospective cohort study using multiple machine learning models. J Orthop Surg Res. 2024;19:575.
13. Choi E, Park D, Son G, et al. Weakly supervised deep learning for diagnosis of multiple vertebral compression fractures in CT. Eur Radiol. 2024;34:3750-60.
14. Chen J, Liu S, Li Y, et al. Deep learning model for automated detection of fresh and old vertebral fractures on thoracolumbar CT. Eur Spine J. 2025;34:1177-86.
15. Lee J, Park H, Yang Z, Woo OH, Kang WY, Kim JH. Improved detection accuracy of chronic vertebral compression fractures by integrating height loss ratio and deep learning approaches. Diagnostics (Basel). 2024;14:2477.
16. Biamonte E, Levi R, Carrone F, et al. Artificial intelligence-based radiomics on computed tomography of lumbar spine in subjects with fragility vertebral fractures. J Endocrinol Invest. 2022;45:2007-17.
17. Dong ST, Zhu J, Yang H, Huang G, Zhao C, Yuan B. Development and internal validation of supervised machine learning algorithm for predicting the risk of recollapse following minimally invasive kyphoplasty in osteoporotic vertebral compression fractures. Front Public Health. 2022;10:874672.
18. Wu H, Li C, Song J, Zhou J. Developing predictive models for residual back pain after percutaneous vertebral augmentation treatment for osteoporotic thoracolumbar compression fractures based on machine learning technique. J Orthop Surg Res. 2024;19:803.
19. Singh M, Tripathi U, Patel KK, Mohit K, Pathak S. An efficient deep learning based approach for automated identification of cervical vertebrae fracture as a clinical support aid. Sci Rep. 2025;15:25651.
20. Inigo B, Shen Y, Killeen BD, et al. An intrinsically explainable approach to detecting vertebral compression fractures in CT scans via neurosymbolic modeling. Proc SPIE Int Soc Opt Eng. 2025;13406:134062R.
21. Kazerouni A, Aghdam EK, Heidari M, et al. Diffusion models in medical imaging: a comprehensive survey. Med Image Anal. 2023;88:102846.
22. Tahir A, Saadia A, Khan K, Gul A, Qahmash A, Akram RN. Enhancing diagnosis: ensemble deep-learning model for fracture detection using X-ray images. Clin Radiol. 2024;79:e1394-402.

DOI: 10.4274/forbes.galenos.2026.52028
Forbes J Med 2026;7:128-39

Original Article

Bridging Research and Innovation in Endoscopic Submucosal Dissection: A Publication-Patent Analysis Revealing Systematic Clinical-Technical Divergence

Hüseyin DEMİR[1], Serhat BURMAOĞLU[2], Tuncay AKAR[3], Adnan YAMANOĞLU[4], Sezgin VATANSEVER[5]

[1]İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Health Management, İzmir, Türkiye
[2]İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Data Science and Analytics, İzmir, Türkiye
[3]Muğla Sıtkı Koçman University, Institute of Health Sciences, Department of Health Management, Muğla, Türkiye
[4]İzmir Katip Çelebi University Faculty of Medicine, Division of Internal Medical Sciences, Department of Emergency Medicine, İzmir, Türkiye
[5]İzmir Katip Çelebi University Faculty of Medicine, Department of Internal Medical Sciences, İzmir, Türkiye

Cite as: Demir H, Burmaoğlu S, Akar T, Yamanoğlu A, Vatansever S. Bridging research and innovation in endoscopic submucosal dissection: a publication-patent analysis revealing systematic clinical-technical divergence. Forbes J Med 2026;7:128-39

ABSTRACT

Objective: Clinical research and technological innovation are assumed to evolve in tandem through market-mediated knowledge diffusion, yet systematic evidence for their alignment in medical device domains remains limited. This study examines the relationship between the production of scientific evidence and the commercial development of devices for endoscopic submucosal dissection, a minimally invasive treatment for gastrointestinal neoplasms, which has experienced uneven global diffusion.

Methods: We employed a dual-corpus computational approach to analyze a stratified sample of 2,470 publications [60% highly cited, 20% recent (2023-2025), and 20% randomly selected] and a corpus of 4,526 device patents (2010-2025), using Word2Vec embeddings and Latent Dirichlet Allocation to identify thematic structures. Cross-corpus semantic similarity measured the alignment between clinical research priorities and commercial development trajectories. Results indicate a substantial publication-patent divergence (divergence score =0.792), with no topic pair exceeding 0.5 in semantic similarity (an original, study-specific metric that has not yet been externally validated).

Results: Most publication topics (88.9%) lack corresponding patent activity, while most patent topics (83.3%) develop independently of academic discourse. Seven innovation gaps emerged in which well-documented clinical complications lack technological solutions, particularly in the management of delayed complications, the prevention of strictures, and the control of pain. Temporal analysis showed an inverse relationship between patenting and publication activity (r=-0.949, lag=-5 years), but this finding should be interpreted cautiously given the small number of annual observations (n=16) and the absence of multiple-lag correction; Granger causality tests indicated no significant bidirectional influence. Only 16.1% of patents demonstrate a high citation impact; these are predominantly in materials science.

Conclusion: The findings challenge assumptions of efficient knowledge diffusion in medical device markets, revealing parallel knowledge streams with weak connectivity. The dual-corpus framework offers a generalizable methodology for assessing research-innovation integration across medical specialties.

Keywords: Endoscopic submucosal dissection, knowledge diffusion, patent-publication analysis, medical device innovation, clinical-technological alignment

Received: 24.07.2026
Accepted: 16.09.2026
Publication Date: 02.10.2026

Corresponding Author:
Hüseyin DEMİR
İzmir Katip Çelebi University Faculty of Economics and Administrative Sciences, Department of Health Management, İzmir, Türkiye
huseyin.demir@ikc.edu.tr
ORCID ID: 0000-0002-8990-7228

INTRODUCTION

Endoscopic submucosal dissection (ESD), developed in Japan in the late 1990s, represents a paradigm shift in minimally invasive treatment of gastrointestinal neoplasms.[1,2] The technique overcomes limitations of endoscopic mucosal resection, particularly the piecemeal resection required for lesions exceeding 20 mm and the associated high recurrence rates.[3-5] ESD enables en bloc resection of large superficial lesions while preserving organ function, offering complete resection rates exceeding 90% for early gastric neoplasms, with 5-year survival approaching those of surgical resection.[6-8] Despite growing acceptance as a standard treatment for superficial gastric neoplasms, esophageal squamous cell carcinoma, and select colorectal lesions, the diffusion of ESD beyond Japan has been uneven.[9-11]

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

This heterogeneity reflects patterns documented in knowledge diffusion theory: geographical localization of tacit expertise, knowledge filters impeding commercialization, and institutional mechanisms constraining technology transfer.[12,13] ESD's technical complexity generates substantial tacit knowledge that resists codification and is transmitted primarily through hands-on training.[14] Technical demands are reflected in complication rates of 3.5% for gastric, 3.3% for esophageal, and 4.6% for colorectal procedures, with operator experience critically determining safety.[15,16] In parallel, a robust commercial ecosystem has emerged, featuring sophisticated instrumentation: electrosurgical knives, injection systems, and hemostatic devices.[17] Medical device innovation, characterized by iterative cycles between clinical use and technological refinement and involving complex interactions among physicians, engineers, regulatory bodies, and manufacturers, differs fundamentally from pharmaceutical development.[18-20] This co-evolutionary dynamic suggests that alignment between clinical evidence and commercial device development should emerge naturally, yet evidence from medical device markets reveals frequent misalignment.[21,22]

Integrated patent-publication databases now provide an unprecedented opportunity to systematically examine clinical-technical alignment. Combined with natural language processing and unsupervised machine learning, these databases enable innovation intelligence that transcends single-corpus limitations.[23] Despite a rich literature on ESD clinical outcomes, comprehensive macro-level analysis connecting publications to patents remains limited. This study addresses clinical-technical integration in ESD through three research questions, the first of which is: (RQ1) To what extent do clinical research themes align with commercial device development priorities? (RQ2) Which specific clinical complications lack corresponding technological solutions? (RQ3) What temporal relationship exists between publication activity and patenting trends, and does evidence suggest bidirectional knowledge transfer? We employ a dual-corpus computational approach, integrating 2,470 publications and 4,526 patents, to quantify alignment and identify structural knowledge filters impeding translation, thereby contributing a generalizable framework with implications for clinicians, training programs, research funders, device manufacturers, and policymakers.

METHODS

Data Acquisition

Publication and patent data were retrieved from Lens.org.[24] The publication dataset encompassed 17,830 peer-reviewed articles identified through queries combining "endoscopic submucosal dissection," "ESD procedure," and "submucosal dissection technique" with organ-specific terms (2010-2025). The patent dataset comprised 5,096 utility patents identified using technology-focused queries. A stratified sampling strategy was implemented to select publications: 60% highly cited, 20% recent (2023-2025), and 20% randomly selected publications, yielding 2,470 documents. "Highly cited" was operationalized as the total citation count within the top decile for its publication year according to Lens.org indexing. The exact numbers of records screened into and retained within each of the three tiers, by year, and the procedure used to resolve overlap between tiers (e.g., when a record qualified as both highly cited and recent) were reported. The tiered design was informed by general bibliometric sampling conventions; however, it did not in itself validate the specific sampling design. A dedicated methodological reference on stratified or citation-tiered sampling in bibliometric research was also provided.[25]

The exclusion workflow was documented by reporting the number of records excluded at each stage and the corresponding reasons for exclusion, including duplicate records, non-English-language records, records with unavailable abstracts, records with fewer than 20 tokens after preprocessing, and records of non-primary-research document types. For patent records, the unit of analysis was explicitly defined as granted patents. To avoid multiple counting of the same invention across jurisdictions, patent records were deduplicated, collapsing records to Lens.org simple-family identifiers where applicable.

Text Preprocessing

All documents underwent a comprehensive NLP pipeline that integrated the title and abstract fields. Research demonstrates that abstracts capture 80-90% of semantic information for topic modeling in medical domains.[26] The preprocessing sequence included conversion to lowercase, removal of URLs and emails, removal of numeric characters, tokenization, removal of stopwords, and WordNet lemmatization. Documents with fewer than 20 tokens were excluded.[27-29]

Semantic Clustering and Topic Modeling

Word2Vec models were trained separately for publications and patents using the CBOW architecture (100-dimensional vectors, context window of 5, minimum word frequency of 5).[30] K-means clustering was applied to document embeddings, and the optimal number of clusters was determined using the elbow method, silhouette coefficients, and Davies and Bouldin[32] indices.[31] Visualization was performed using t-SNE (perplexity =30, 1,000 iterations).[33] Probabilistic topic modeling employed latent Dirichlet allocation (LDA), with the optimal number of topics determined by maximization of the CV coherence score across k=2-12 (publications) and k=2-10 (patents).[34,35] Final models: 9 publication topics and 6 patent topics (15 passes, 400 iterations).

Innovation Divergence Quantification

The two LDA models were fitted independently on their respective corpora (publications, patents), each producing a topic-term probability distribution over its own vocabulary. A common vocabulary was constructed (e.g., as the union or intersection

of terms retained after preprocessing in both corpora, or as the top-N TF-weighted terms per topic), and each topic's term-probability vector was renormalized to this vocabulary before computing cosine similarity, producing a 9×6 (publication-topic × patent-topic) similarity matrix, i.e., 54 pairwise comparisons. Topic labels (e.g., "procedural technique," "electrosurgical tools") were assigned by independent review of the top 15 terms per topic and were applied separately to the publication and patent models, since the two were not jointly estimated.

Cross-corpus alignment was assessed by extracting topic-term probability matrices from both LDA models, identifying shared vocabulary (typically 40-60% overlap), and computing cosine similarity for all publication-patent topic pairs. Topic pairs with similarity ≥0.5 were classified as strongly aligned; publication topics with maximum similarity <0.3 were designated "orphan" topics. The divergence score was calculated as: Divergence = 1-(mean maximum similarity across all topics).

Classification into "strongly aligned" versus "orphan" categories used fixed descriptive similarity thresholds (≥0.5, <0.3) rather than a per-pair null-hypothesis significance test, so a Bonferroni correction does not directly apply to these categorical cut-offs. When inferential statistics (e.g., permutation-based p values) are also reported for any of the 54 topic-pair comparisons, a multiple-comparison correction should be applied, and the corrected results should reported alongside the raw similarity values.

Temporal Dynamics

Annual time series spanning 2010-2025 were constructed for both corpora. Cross-correlation analysis examined lead-lag relationships across lags of -10 to +10 years on the detrended series. Granger causality testing employed a vector autoregression with lags of 1-3 years (α=0.05). Because the underlying series comprise only 16 annual observations (2010-2025), the VAR and Granger causality estimates are based on a sample that is too small for reliable inference; this constraint constitutes a major limitation, and the results should be interpreted as exploratory rather than confirmatory. Granger causality was evaluated using the sum of squared residuals F-test applied to the detrended publication and patent series for lags 1 through 3. Because this implementation tests each specified lag independently rather than selecting a single lag via an information criterion and does not perform an internal stationarity test or post-estimation residual diagnostics, no formal stationarity or residual diagnostic tests were conducted on the underlying series in the present analysis. This is acknowledged as an additional methodological limitation, and we recommend that a formal stationarity assessment and residual diagnostics be incorporated into future work to strengthen the reliability of causal inference. The effective number of observations available at each of the three tested lags is presented, together with 95% confidence intervals for the reported correlation and F-statistics.[36] Compound annual growth rates (CAGR) were calculated for both time series.

Complication-Device Mapping

A complication taxonomy encompassing five categories (bleeding, perforation, stricture, infection, pain) and encompassing a device taxonomy (dissection knives, hemostasis devices, injection systems, closure devices, cautery instruments) were established. Keyword-based classification identified mentions of complications and of device types in abstracts. A complication-device solution matrix was used to cross-tabulate these categories and reveal innovation gaps. Technology maturity was assessed by patent age (<5, 5-10, >10 years) and forward citation counts (low <5, moderate 5-10, high ≥10). All analyses were performed using Python 3.12 with a fixed random seed [42] for reproducibility. Informed consent was not required for this study, as it involved the bibliometric analysis of publicly available bibliographic data and did not include human participants or the collection or processing of personally identifiable information. This study was based exclusively on publicly available publication data retrieved from established bibliometric databases. No human participants were involved, and the study did not entail primary data collection or any form of intervention. As the data analyzed were already publicly accessible and did not contain identifiable personal information or sensitive data, institutional review board or ethics committee approval was not required for this bibliometric analysis.

Findings

Semantic Structure and Thematic Organization

Detailed findings regarding the semantic structure and thematic organization are presented in Figure 1a-d. Word2Vec clustering identified k=2 as the optimal number of clusters for publications (silhouette =0.188) and k=3 as the optimal number of clusters for patents (silhouette =0.157). Publication Cluster 0 (51.3%) emphasized procedural-technical content; Cluster 1 (48.7%) focused on oncological outcomes. Patent clustering demonstrated pronounced segregation: Cluster 0 (38.2%, mechanical innovations), Cluster 1 (33.5%, injection/lifting technologies), and Cluster 2 (28.3%, materials science/coatings). When visualized in the same semantic space, these clusters show minimal spatial overlap, providing early evidence of thematic divergence. The silhouette coefficients obtained for both clustering solutions were low (0.188 for publications; 0.157 for patents), indicating weak cluster separation in the Word2Vec embedding space. These Word2Vec + K-means clusters, therefore, provide only preliminary, low-confidence evidence of thematic structure and should be interpreted as complementary to, and subordinate to, the LDA topic model results, which form the basis of the main divergence analysis.

LDA coherence optimization identified 9 topics in publications (coherence =0.465, k=9) and 6 topics in patents (coherence =0.463, k=6). Figure 2a-d provide a comprehensive overview of the key topics and their associated word clouds. This dimensional asymmetry—nine clinical topics mapping onto six patent topics—represents a structural obstacle to knowledge transfer.

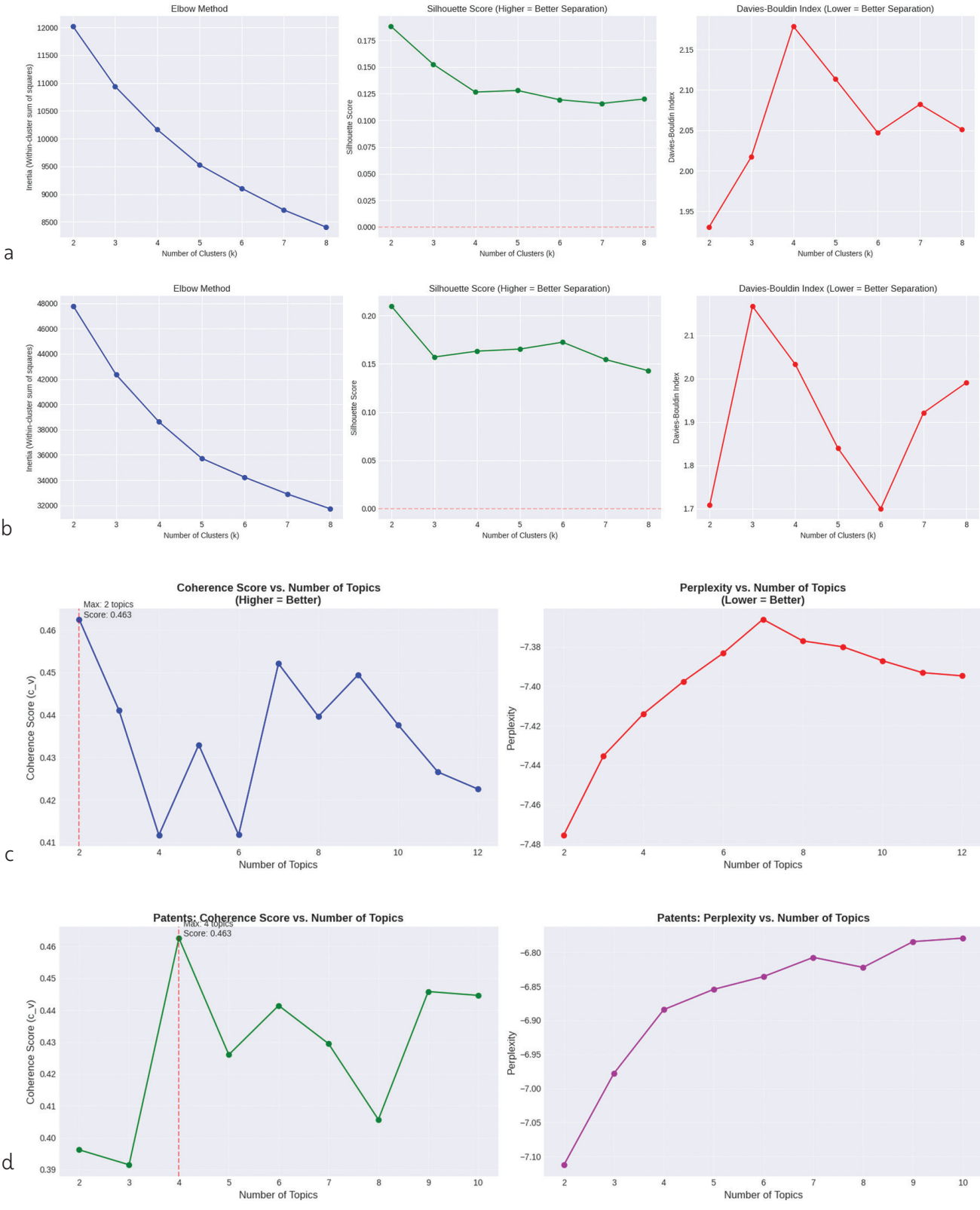


Figure 1. Cluster optimization metrics for publications (a) and patents (b); LDA topic coherence optimization curves for publications (c) and patents (d)

LDA: Latent Dirichlet allocation

Publication topics spanned procedural technique (T0), patient outcomes (T1), cancer characteristics (T2), technique refinement (T3), adverse events (T4), comparative studies (T5), training/competency (T6), anatomical applications (T7), and classification systems (T8). Patent topics covered: device mechanics (T0), electrosurgical tools (T1), material compositions (T2), injection systems (T3), imaging/navigation (T4), and closure/sealing (T5).

Innovation Landscape Divergence (RQ1)

The publication-patent alignment matrix quantified semantic similarity across all 54 possible topic pairs. Figure 3a-e provide a detailed overview of the innovation landscape divergence. No topic pair achieved strong alignment (≥0.5), and only two achieved moderate alignment: publication T3 (technique refinement) aligned with patent T1 (electrosurgical tools) at 0.316—the maximum observed—while publication T4 (adverse events) aligned with patent T5 (closure devices) at 0.289. Eight of nine publication topics (88.9%) showed no meaningful alignment (score <0.3) with any patent topic, including patient outcomes (maximum 0.18), training (maximum 0.19), and classification systems (maximum 0.17). Reciprocally, five of the six patent topics (83.3%) develop independently of academic research. The divergence score of 0.792 [Divergence = 1-(0.316+0.289 + … + 0.282)/15=1-0.208] is an original metric introduced in this study and has not yet been validated or benchmarked against other studies or device domains, and is descriptively consistent with a substantial clinical-technical misalignment.

Temporal Innovation Dynamics (RQ3)

Figure 4a-c illustrate the temporal dynamics of innovation in detail. ESD publications increased from 89 (2010) to a peak of 241 (2015) and then declined sharply (CAGR =-21.5%), reaching 82 by 2024—a 66% reduction from the peak. In contrast, patent filings grew steadily (CAGR =+8.9%) from 138 (2010) to 479 (2025) without reaching a plateau. The crossover point occurred in 2015; by 2025, patents outnumbered publications by 7.4:1. This inverse relationship persisted across all individual topics: all nine publication topics declined after the 2015-2018 peaks, whereas all six patent topics maintained robust growth through 2025.

Cross-correlation analysis revealed the strongest negative correlation at lag = -5 years ($r=-0.949$, $p<0.001$). Given that only 16 annual observations underlie this analysis and that 21 lags (-10 to +10 years) were examined, this correlation should be interpreted with caution: it does not, by itself, establish that patent growth temporally precedes publication decline, and no correction for the 21 lag-wise comparisons has yet been applied. Because the corresponding Granger causality tests were not significant in both directions, this cross-correlation result is presented as a descriptive, hypothesis-generating association rather than evidence of a lead-lag causal relationship. A weak positive correlation at lag =+3 years ($r=0.187$, $p=0.342$) and a near-zero contemporaneous correlation ($r=-0.143$, $p=0.476$) indicate that publication trends do not predict patent activity. Granger causality tests yielded nonsignificant results in both directions (patents→publications: $F=1.83$, $p=0.17$; publications→patents: $F=1.24$, $p=0.31$), suggesting a shared response to external confounding factors rather than a direct causal influence.

Complication Management Technology Landscape (RQ2)

Figure 5a-e provide a detailed overview of the complication management technology landscape. The most frequently mentioned complications were bleeding (486, 19.7%) and perforation (469, 19.0%), followed by stricture (270, 10.9%), infection (194, 7.9%), and pain (64, 2.6%). Reported values reflect the number of publications in which a given complication keyword was detected. Publication-to-patent ratios revealed striking mismatches (publication mentions per corresponding patent; e.g., bleeding: 486 publication mentions vs. 136 patents ≈3.6 publications per patent): bleeding 3.6:1, perforation 6.2:1, stricture 2.8:1, infection 2.2:1, and pain 9.1:1. Closure devices were most prevalent (n=508), followed by dissection knives (n=452), injection systems (n=423), hemostasis devices (n=222), and cautery instruments (n=20).

Seven innovation gaps were identified from the complication-device solution matrix:

Gap 1-Delayed Bleeding Management (Critical): The 12 studies targeting delayed hemorrhage lack integration with clinical prediction tools that could guide prophylactic interventions.

Gap 2-Stricture Prevention in Circumferential Resection (Critical): Eighty-seven of 98 stricture patients focus on post-stricture dilation rather than prevention. Tissue engineering approaches remain absent from the patent landscape.

Gap 3-Pain Control for Extended Procedures (Moderate-High): Only 7 patents address sedation monitoring without targeting pain sources; 64 clinical mentions produce a 9.1:1 gap ratio.

Gap 4-Submucosal Fibrosis Navigation (High): Twenty-three tissue characterization patents detect fibrosis but provide no actionable dissection solutions.

Gap 5-Perforation Risk Prediction (Moderate): Seventy-six patents address closure; only 8 address prediction. Integrated predictive models guiding real-time decision-making are absent.

Gap 6-Mucosal Defect Size Estimation (Low-Moderate): Twenty measurement patents lack decision-support algorithm integration.

Gap 7-Post-ESD Infection Management (Low): Despite the second-best ratio (2.2:1), innovations targeting specific ESD infection pathways remain absent.

These gaps reveal a systematic pattern: commercial innovation gravitates toward acute, intra-procedural, device-amenable problems while neglecting delayed complications, prevention (rather than treatment), and patient-centered endpoints.

A technology-maturity assessment revealed that 46.7% of patents were recent (<5 years), 35.9% were established (5-10 years), and 17.4% were mature (>10 years), indicating active innovation rather than market consolidation. Paradoxically, 74.4% demonstrated low citation impact (<5 forward citations), with only 16.1% achieving high impact (≥10 citations). High-impact patents were concentrated in materials science (48%), electrosurgical energy (31%), and mechanical innovations (21%), suggesting clinical impact depends more on enabling technology platforms than on individual device designs.

Integration and Synthesis

Three critical patterns emerge: simultaneous high divergence (0.792), declining publications (-21.5% CAGR), and rising patents (+8.9% CAGR) reveal a fundamental disconnection, wherein commercial innovation proceeds independently of academic research—these opposite-directional trends contradict knowledge diffusion theory's prediction of increasing convergence in mature domains. Second, even the most-studied complications remain inadequately addressed by technology (ratios of 3.6:1 for bleeding, 6.2:1 for perforation), challenging the assumption that market forces efficiently translate documented clinical needs into innovation opportunities. Third, the predominance of recent, low-impact patents (46.7% are <5 years old; 74.4% have <5 citations) indicates incremental rather than disruptive advancement.

The ESD innovation ecosystem suffers from three interconnected pathologies: misaligned priorities, insufficient translation, and high noise-to-signal ratio, which warrant strategic intervention rather than reliance on market forces alone.

DISCUSSION

This study provides the first comprehensive integration of publication and patent data to characterize the clinical and technical innovation landscape in ESD. The near-complete absence of strong alignment between publication and patent topics (zero pairs with >0.5 similarity among 54 combinations) provides empirical confirmation of the "knowledge filter": institutional and informational barriers separating clinical researchers from industrial engineers.[37,38]

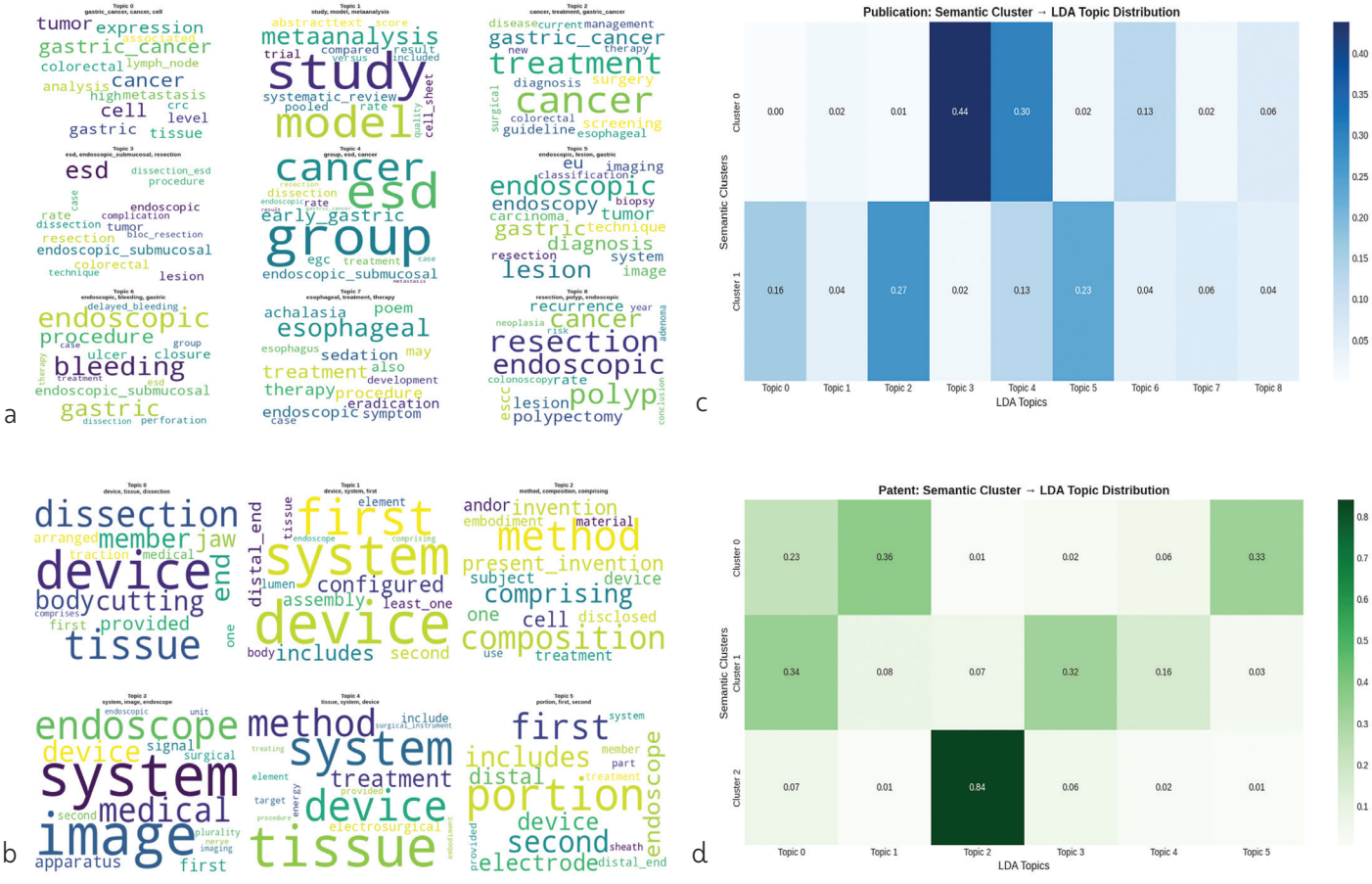


Figure 2. Topic word clouds for publication (a) and patent topics (b), Cluster-topic distribution heatmaps for publications (c) and patents (d)

LDA: Latent Dirichlet allocation

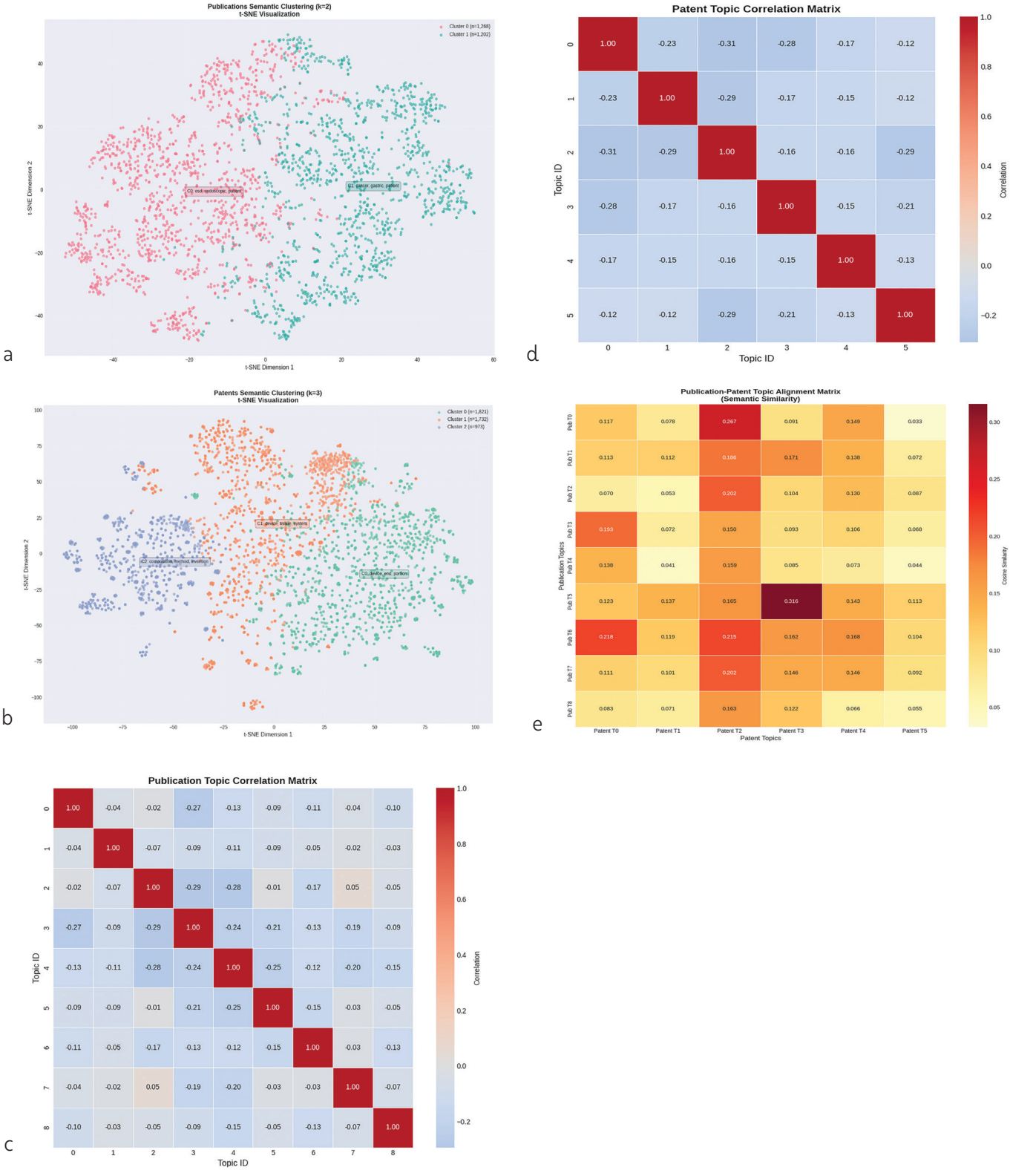
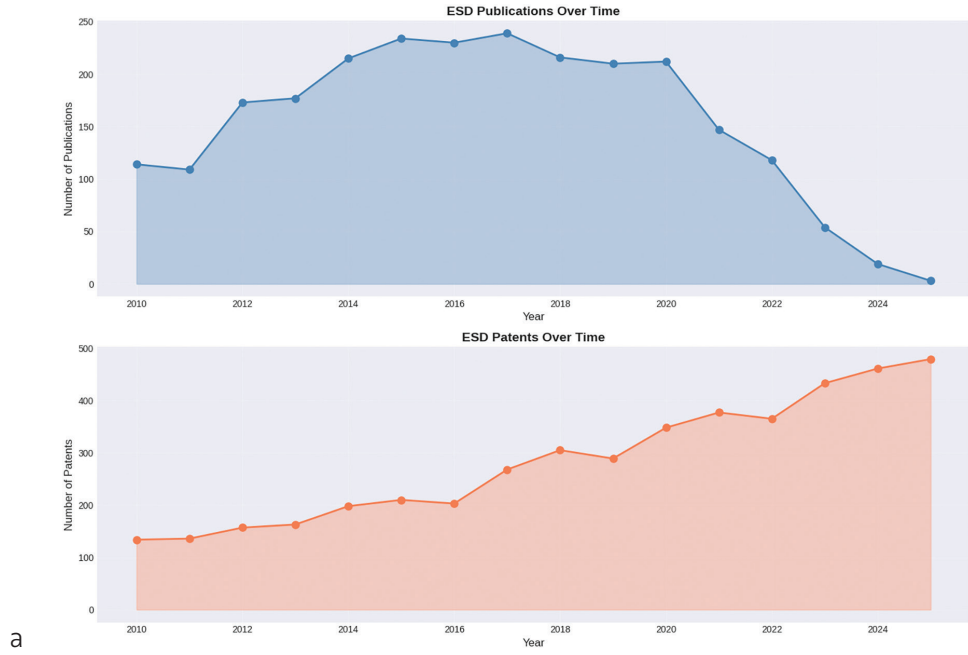
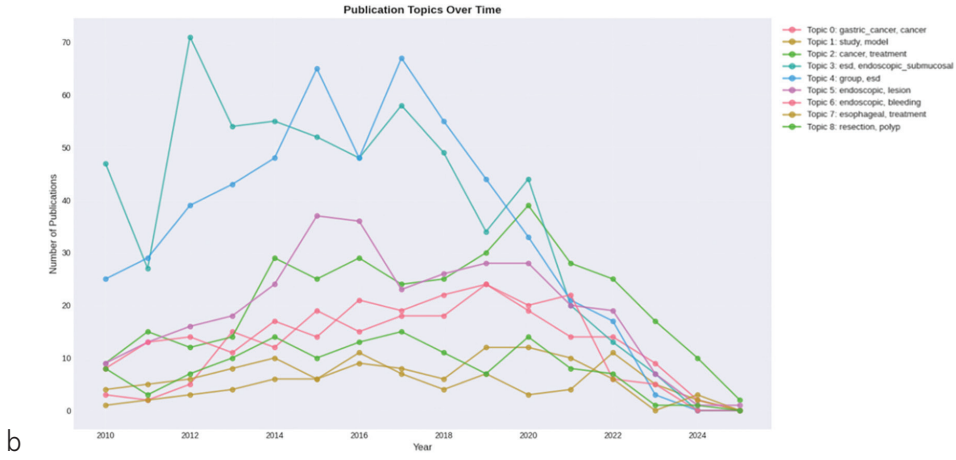


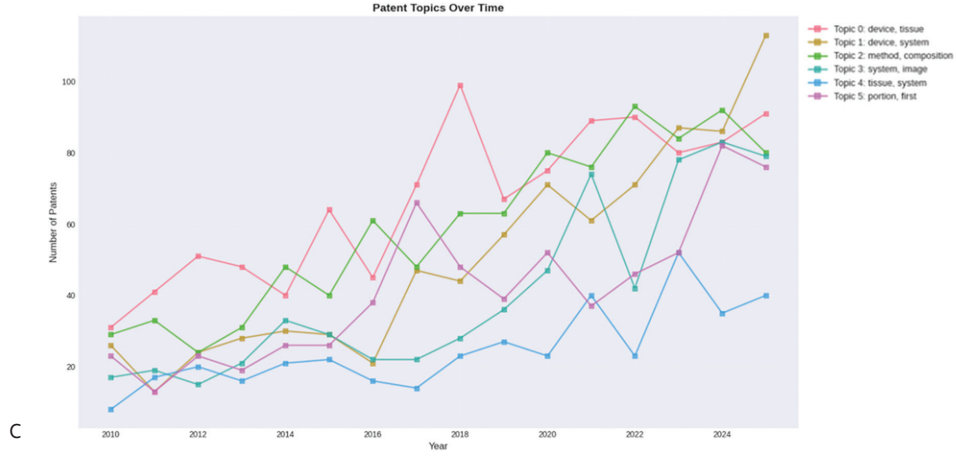
Figure 3. Within-corpus topic correlation matrices for publications (a) and patents (b), t-SNE visualization of semantic clustering for publications (c) and patents (d), and Publication-patent topic alignment matrix (e)



a



b



c

Figure 4. Temporal trends (a), individual topic evolution over time for publications (b) and patents (c)
ESD: Endoscopic submucosal dissection

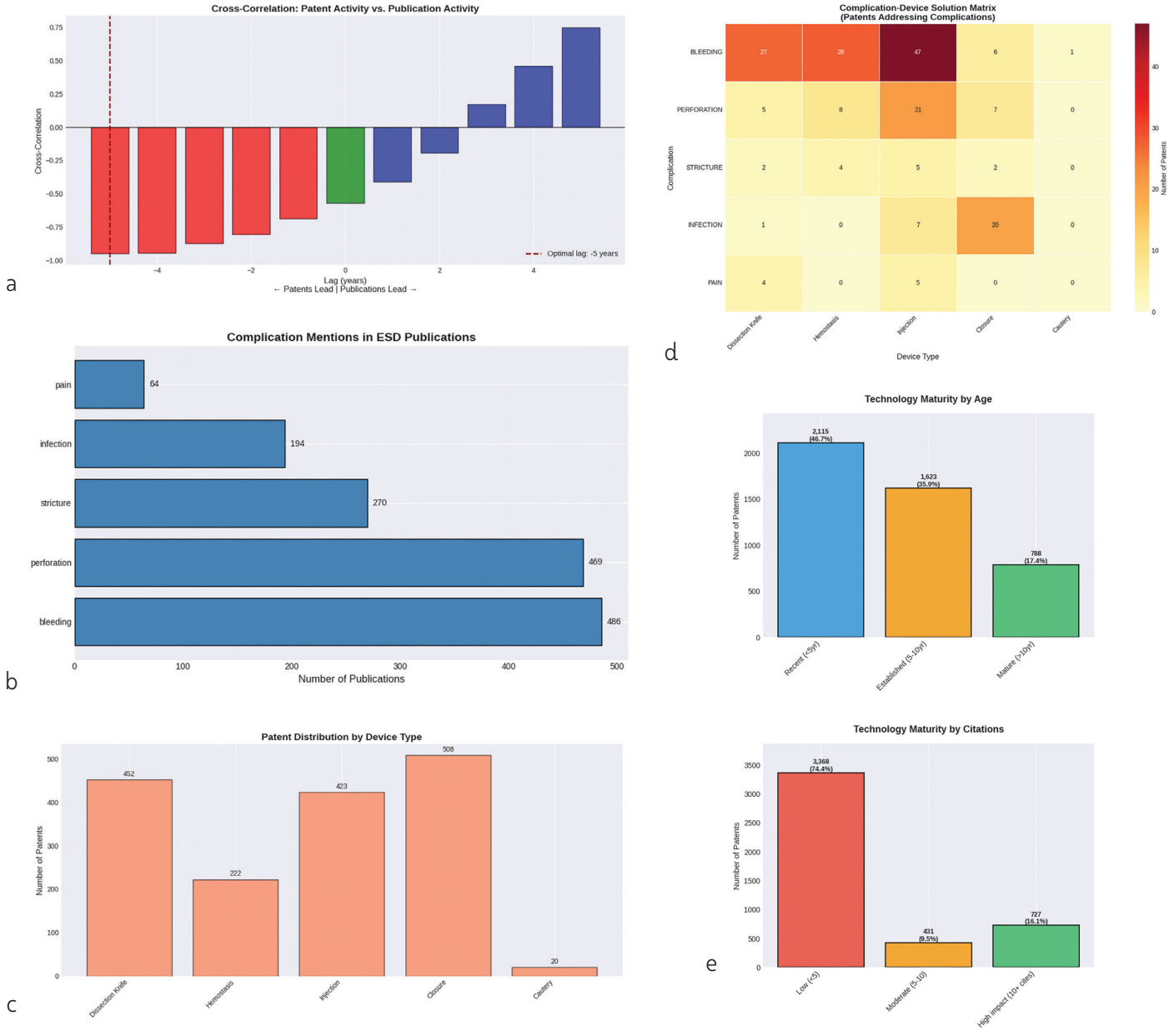


Figure 5. Cross-correlation (a) and complication mentions in ESD Publications (b), patent distribution across five device categories (c), complication-device solution matrix (d), and technology maturity metrics (e)

ESD: Endoscopic submucosal dissection

Several mechanisms explain this divergence: a temporal mismatch between publication cycles and patent-filing strategies; divergent incentive structures (clinical novelty vs. competitive advantage in device specifications); proprietary commercial development that renders substantial innovation invisible to the published literature; and a geographical concentration of ESD expertise in Japan during 2010-2018, resulting in parallel trajectories.[13,39,40]

Innovation Gaps and Knowledge Filter Mechanisms

The complication-solution matrix reveals heterogeneous technology availability, consistent with uneven absorptive capacity. For bleeding (486 publication mentions, 136 patents, 3.6:1 ratio), the technological emphasis on prevention rather than reactive management may reflect industry risk assessment, although both needs exist. Perforation exhibits more severe gaps (6.2:1): the

predominance of injection-based prevention over closure devices permits two interpretations: either effective knowledge transfer that prioritizes prevention or a failure to develop post-perforation repair technologies. The seven complication-device combinations exhibiting zero patent coverage represent a complete blockage of the knowledge filter. Pain management (9.1:1 ratio) represents the clearest gap: procedural pain remains outside the absorptive capacities of academic and commercial ecosystems.

Temporal Dynamics and Ecosystem Maturation

The inverse temporal relationship ($r=-0.949$ at -5 year lag), which should be interpreted cautiously given the small number of annual observations underlying it, contradicts conventional expectations. When combined with non-significant Granger causality (both directions $p>0.05$), these results suggest a shared response to external factors rather than a direct causal influence. The decline in publications following the 2015 peak may reflect the maturation of clinical evidence: as ESD's efficacy became well-established, diminishing returns reduced incentives to publish while commercial pressures intensified, consistent with innovation lifecycle theory. Changing funding priorities, global diffusion dynamics, and regulatory changes represent additional potential drivers. The convergence of all nine publication topics toward near-zero activity by 2024-2025 suggests systemic factors rather than topic-specific dynamics.

Technology Maturity and Innovation Quality

The paradox of quantitatively active yet qualitatively incremental innovation (46.7% recent patents and 74.4% low-citation) is consistent with findings that healthcare innovation evolves through asynchronous, institutionally fragmented pathways. Technical constraints of endoscopic platforms, regulatory incentives favoring 510(k) clearance for substantially equivalent devices, and reimbursement structures that do not differentiate based on device sophistication collectively create a bounded innovation space. The concentration of high-impact patents in materials science (48%) rather than in mechanical innovations (21%) suggests that novel chemistry is a more influential approach, aligning with knowledge diffusion patterns where breakthrough innovations emerge from cross-disciplinary transfer. The 510(k) pathway, reimbursement structure, and cross-disciplinary transfer mechanisms proposed above are offered as plausible explanatory hypotheses consistent with the observed patterns. This study measures textual similarity and temporal co-variation between publication and patent corpora and does not directly measure regulatory pathway choice, reimbursement policy, technology adoption, or clinical benefit; therefore, these mechanisms have not been directly tested here and should be confirmed with regulatory and health-economics data in future work.

Implications for Practice and Policy

In clinical practice, the technology inventory enables evidence-based device selection and identifies capability gaps; when complications lacking robust solutions—particularly delayed bleeding and stricture formation—clinicians should emphasize refinement of operator technique over device-centric approaches. For training programs, the 46.7% proportion of recent patents suggests curricula that emphasize fundamental principles over device-specific techniques, while the 17.4% of mature patents identify stable foundations for core skills. Regarding research funding, targeted translational investment could bridge gaps in stricture prevention and pain management through institutional mechanisms that facilitate academic-industrial knowledge transfer. From an innovation policy perspective, the predominance of incremental innovation suggests that current regulatory and reimbursement structures do not sufficiently incentivize high-impact innovation, warranting coordinated action across 510(k) pathways and bundled reimbursement structures.

Study Limitations

Limitations include citation-based sampling, which introduces recency bias for 2023-2025 publications; reliance on abstracts rather than full-text, although prior research demonstrates 80-90% semantic information capture; an English-language patent-family focus that may underrepresent Japanese domestic filings, mitigated by Lens.org's PCT coverage; and a 16-year temporal scope potentially insufficient for detecting longer innovation cycles. Additional limitations identified during peer review include: the Granger causality and vector autoregression analyses rest on only 16 annual observations (2010-2025), which is a very small sample for reliable VAR/Granger estimation; therefore, temporal-causal conclusions should be treated as exploratory; the cross-correlation analysis examined 21 lags on a 16-point series without a formal multiple-comparison correction; therefore, the lag =-5-year result requires cautious interpretation and, ideally, a corrected significance threshold; the "divergence score" is a study-specific metric that has not been externally validated or benchmarked against other medical device domains; complication and device classification relied on keyword matching whose precision and recall against manually coded labels has not yet been independently verified; this study measures textual similarity and temporal co-variation only, and does not directly measure technology adoption, clinical benefit, regulatory pathway choice [e.g., 510(k) status], reimbursement outcomes, or knowledge transfer; therefore, causal or mechanistic interpretations of the observed associations remain hypotheses for future testing; and the Word2Vec + K-means clustering solutions showed weak cluster separation (silhouette =0.188 for publications, 0.157 for patents) and should be regarded as a preliminary, lower-confidence complement to the LDA-based topic analysis rather than independent confirmation of it. Future directions include full-text analysis, author-inventor linkage, longitudinal device family tracking, comparative specialty analysis, patient outcome linkage, and geographic comparative analysis. Future work should include a formal validation study of the keyword-based

classification pipeline and, where feasible, present the annual publication and patent counts, topic labels and key terms, the full 9×6 similarity matrix, complication-device cell counts, and patent maturity figures as supplementary tables rather than solely as figure panels, to improve readability and reproducibility.

CONCLUSION

This study demonstrates that clinical research and commercial innovation in ESD operate as parallel, largely independent streams. The 0.792 divergence score, zero strongly-aligned topic pairs, seven innovation gaps, and inverse temporal dynamics collectively reveal an innovation ecosystem requiring deliberate integration mechanisms rather than relying on market-mediated knowledge diffusion. The predominance of incremental innovation (74.4% low-impact patents), despite active filing rates (46.7% recent), suggests that current incentive structures favor patent proliferation over transformative advancement.

For stakeholders, this analysis provides actionable intelligence: clinicians receive evidence-based device inventories; training programs gain guidance on application priorities; research funders identify gaps for targeted investment; device manufacturers understand competitive landscapes and unmet needs; and regulators recognize where approval pathways may insufficiently incentivize needed innovation. The integrated publication-patent methodology offers a generalizable framework for innovation intelligence across procedure-based medical specialties. The path forward requires coordinated action: research funding mechanisms that incentivize technology evaluation; regulatory pathways that balance evidence requirements with innovation incentives; reimbursement structures that reward superior outcomes; and institutional models that facilitate author-inventor overlap.

Ethics

Ethics Committee Approval: This study was based exclusively on publicly available publication data retrieved from established bibliometric databases. No human participants were involved, and the study did not entail primary data collection or any form of intervention. As the data analyzed were already publicly accessible and did not contain identifiable personal information or sensitive data, institutional review board or ethics committee approval was not required for this bibliometric analysis.

Informed Consent: Informed consent was not required for this study, as it involved the bibliometric analysis of publicly available bibliographic data and did not include human participants or the collection or processing of personally identifiable information.

Footnotes

Authorship Contributions

Concept: H.D., S.B., T.A., A.Y., S.V., Design: H.D., S.B., T.A., A.Y., S.V., Data Collection or Processing: H.D., S.B., T.A., A.Y., S.V., Analysis or Interpretation: H.D., S.B., T.A., A.Y., S.V., Literature Search: H.D., S.B., T.A., A.Y., S.V., Writing: H.D., S.B., T.A., A.Y., S.V.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Fukuzawa M, Gotoda T. History of endoscopic submucosal dissection and role for colorectal ESD. Gastrointest Interv. 2012;1:30-5.
2. Yamamoto H. Endoscopic submucosal dissection--current success and future directions. Nat Rev Gastroenterol Hepatol. 2012;9:519-29.
3. Yamamoto H. Endoscopic submucosal dissection of early cancers and large flat adenomas. Clin Gastroenterol Hepatol. 2005;3:S74-6.
4. Elkholy S, Awad A, Haggag H, Essam K, Yousef K, El Sherbiny M. Different endoscopic submucosal dissection techniques: a tailored approach. VideoGIE. 2024;10:277-84.
5. Bourke MJ, Neuhaus H, Bergman JJ. Endoscopic submucosal dissection: indications and application in western endoscopy practice. Gastroenterology. 2018;154:1887-900.e5.
6. Harlow C, Sivananthan A, Ayaru L, Patel K, Darzi A, Patel N. Endoscopic submucosal dissection: an update on tools and accessories. Ther Adv Gastrointest Endosc. 2020;13:2631774520957220.
7. Lambin T, Rivory J, Wallenhorst T, et al. Endoscopic submucosal dissection: how to be more efficient? Endosc Int Open. 2021;9:E1720-30.
8. Draganov PV, Wang AY, Othman MO, Fukami N. AGA institute clinical practice update: endoscopic submucosal dissection in the United States. Clin Gastroenterol Hepatol. 2019;17:16-25.e1.
9. Matsui N, Akahoshi K, Nakamura K, Ihara E, Kita H. Endoscopic submucosal dissection for removal of superficial gastrointestinal neoplasms: a technical review. World J Gastrointest Endosc. 2012;4:123-36.
10. Mendes RR, Barreiro P, Mascarenhas A, Franco AR, Carvalho L, Chagas C. Endoscopic submucosal dissection for resections larger than 10 cm: outcomes from a portuguese center. GE Port J Gastroenterol. 2023;31:33-40.
11. Kim DH, Jin BC, Oh HH, et al. Outcomes of colorectal endoscopic submucosal dissection according to the size of colorectal neoplasm: a HASID multicenter study. In Vivo. 2024;38:2009-15.
12. Autant-Bernard C, Fadairo M, Massard N. Knowledge diffusion and innovation policies within the European regions. Res Policy. 2013;42:196-210.
13. Bahar D, Hausmann R, Hidalgo CA. Neighbors and the evolution of the comparative advantage of nations. J Int Econ. 2014;92:111-23.
14. Barbour JA, O'Toole P, Suzuki N, Dolwani S. Learning endoscopic submucosal dissection in the UK: barriers, solutions and pathways for training. Frontline Gastroenterol. 2020;12:671-6.
15. Odagiri H, Yasunaga H. Complications following endoscopic submucosal dissection for gastric, esophageal, and colorectal cancer: a review of studies based on nationwide large-scale databases. Ann Transl Med. 2017;5:189.
16. Kuroha M, Shiga H, Kanazawa Y, et al. Factors associated with fibrosis during colorectal endoscopic submucosal dissection: does pretreatment biopsy potentially elicit submucosal fibrosis and affect endoscopic submucosal dissection outcomes? Digestion. 2021;102:590-8.
17. Toyonaga T, Man-i M, East JE, et al. 1,635 Endoscopic submucosal dissection cases in the esophagus, stomach, and colorectum:

complication rates and long-term outcomes. Surg Endosc. 2013;27: 1000-8.

18. Kaplan AV, Baim DS, Smith JJ, et al. Medical device development: from prototype to regulatory approval. Circulation. 2004;109:3068-72.
19. Chatterji AK, Fabrizio KR, Mitchell W, Schulman KA. Physician-industry cooperation in the medical device industry. Health Aff (Millwood). 2008;27:1532-43.
20. Lettl C, Rost K, von Wartburg I. Why are some independent inventors 'heroes' and others 'hobbyists'? The moderating role of technological diversity and specialization. Res Policy. 2009;38:243-54.
21. Cohen D, Billingsley M. Europeans are left to their own devices. BMJ. 2011;342:d2748.
22. Murray F, Stern S. Do formal intellectual property rights hinder the free flow of scientific knowledge? J Econ Behav Organ. 2007;63:648-87.
23. Vajjala S, Majumder B, Gupta A, Surana H. Practical natural language processing. Sebastopol: O'Reilly Media; 2020.
24. The Lens. The Lens: Free & open patent and scholarly literature search [Internet]. Available from: https://about.lens.org/
25. Zupic I, Čater T. Bibliometric methods in management and organization. Organizational Research Methods. 2015;18:429-72.
26. Syed S, Spruit M. Full text or abstract? Examining topic coherence scores using latent Dirichlet allocation. In: 2017 IEEE International Conference on Data Science and Advanced Analytics (DSAA); 2017; Tokyo, Japan. IEEE; 2017. p. 165-74.
27. Bird S, Klein E, Loper E. Natural language processing with python. Sebastopol: O'Reilly Media. 2009.
28. Miller GA. WordNet: a lexical database for English. Commun ACM. 1995;38:39-41.
29. Maier D, Waldherr A, Miltner P, et al. Applying LDA topic modeling in communication research. Commun Methods Meas. 2018;12:93-118.
30. Mikolov T, Chen K, Corrado G, Dean J. Efficient estimation of word representations in vector space. arXiv. 2013.
31. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53-65.
32. Davies DL, Bouldin DW. A cluster separation measure. IEEE Trans Pattern Anal Mach Intell. 1979;1:224-7.
33. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9:2579-605.
34. Blei DM, Ng AY, Jordan MI. Latent Dirichlet allocation. J Mach Learn Res. 2003;3:993-1022.
35. Röder M, Both A, Hinneburg A. Exploring the space of topic coherence measures. WSDM. 2015.
36. Granger CWJ. Investigating causal relations by econometric models and cross-spectral methods. Econometrica. 1969;37:424-38.
37. Braunerhjelm P, Acs S, Audretsch D, Carlsson B. The missing link: knowledge diffusion and entrepreneurship in endogenous growth. Small Bus Econ. 2010;34:105-25.
38. Audretsch DB, Bönte W, Keilbach M. Entrepreneurship capital and its impact on knowledge diffusion. J Bus Venturing. 2008;23:687-98.
39. Bacchiocchi E, Montobbio F. Knowledge diffusion from university and public research. J Technol Transfer. 2009;34:169-81.
40. Baruffaldi SH, Simeth M. Patents and knowledge diffusion: the effect of early disclosure. Res Policy. 2020;49:103927.

DOI: 10.4274/forbes.galenos.2026.68916
Forbes J Med 2026;7:140-6

Original Article

Non-steroidal Anti-inflammatory Drug Hypersensitivity in Children: From Urticaria to NSAID-Exacerbated Respiratory Disease

Özgen SOYÖZ[1], Figen ÇELEBİ ÇELİK[2], Selime ÖZEN BÖLÜK[2], İlke TAŞKIRDI[2], Mehmet Şirin KAYA[2], İdil AKAY HACI[2], Canan Şule KARKINER[2], Özlem SANCAKLI[2], Demet CAN[2]

[1]Bakırçay University Çiğli Training and Research Hospital, Department of Pediatric Allergy and Immunology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital, Clinic of Pediatric Allergy and Immunology, İzmir, Türkiye

Cite as: Soyöz Ö, Çelebi Çelik F, Özen Bölük S, et al. Non-steroidal anti-inflammatory drug hypersensitivity in children: from urticaria to NSAID-exacerbated respiratory disease. Forbes J Med 2026;7:140-146

ABSTRACT

Objective: Non-steroidal anti-inflammatory drug (NSAID) hypersensitivity is one of the leading causes of drug hypersensitivity in children. However, data on phenotype distribution according to the 2025 World Allergy Organization (WAO) classification and on long-term outcomes remain limited. To describe the clinical characteristics, phenotype distribution, and long-term outcomes of children with NSAID hypersensitivity.

Methods: This retrospective, single-center study included 43 children (1-18 years) diagnosed with NSAID hypersensitivity between 2015 and 2020. Clinical characteristics, culprit drugs, atopic comorbidities, drug provocation test (DPT) results, and follow-up data were reviewed. Hypersensitivity reactions were classified according to the 2025 WAO classification.

Results: Urticaria and/or angioedema were the predominant clinical presentations (83.7%). The most frequently implicated drugs were ibuprofen (65.1%) and paracetamol (41.9%). According to the WAO classification, 46.5% of patients had single NSAID-induced urticaria/angioedema/anaphylaxis, 41.9% had NSAID-induced urticaria/angioedema/anaphylaxis, 7.0% had single NSAID-induced delayed hypersensitivity reactions, and 4.7% had probable NSAID-exacerbated respiratory disease (N-ERD). Nearly half (46.3%) reacted to NSAIDs from multiple chemical classes. DPT was performed in 16 patients (37.2%) either to evaluate the suspected culprit drug or to identify a safe alternative analgesic. During follow-up, 21.4% of patients who used NSAIDs after diagnosis reported tolerance upon culprit-drug re-exposure.

Conclusion: Pediatric NSAID hypersensitivity predominantly presents with cutaneous manifestations, although respiratory phenotypes such as N-ERD also occur. Careful clinical phenotyping, DPT, and identification of safe alternative analgesics are essential for optimal management. Prospective multicenter studies are needed to define the natural history of pediatric NSAID hypersensitivity and the predictors of the development of tolerance.

Keywords: Children, drug hypersensitivity, drug provocation test, nonsteroidal anti-inflammatory drug, NSAID-exacerbated respiratory disease, urticaria

Received: 30.07.2026
Accepted: 16.09.2026
Publication Date: 02.10.2026

Corresponding Author:
Özgen SOYÖZ
Bakırçay University Çiğli Training and Research Hospital, Department of Pediatric Allergy and Immunology, İzmir, Türkiye
ozgensoyoz@yahoo.com
ORCID ID: 0000-0002-4540-352X

INTRODUCTION

Non-steroidal anti-inflammatory drugs (NSAIDs) are widely used in children for fever, pain, and inflammation, and their effects depend largely on blockade of cyclooxygenase (COX). Hypersensitivity to this drug group was long regarded as an adult problem, yet reports in children and adolescents have grown steadily, and NSAIDs now rank alongside antibiotics as the drugs most often implicated in childhood hypersensitivity reactions.[1,2]

NSAIDs are classified into subgroups according to their pharmacological properties, particularly their degree of COX-1 and COX-2 inhibition, and their chemical structures. Based on the underlying pathogenic mechanism, hypersensitivity reactions are broadly categorized as either cross-reactive reactions, which are non-immunological and associated with COX-1 inhibition, or selective reactions, which are immunologically mediated through IgE- or T-cell–dependent mechanisms. The World Allergy Organization (WAO) revised the phenotypic categories of this scheme in 2025 (Table 1).[3,4]

The revised phenotypes are based largely on adult data, and their manifestations in children are much less clear. In pediatric

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

series, skin findings predominate, whereas airway involvement is reported only rarely. Even so, NSAID-exacerbated respiratory disease (N-ERD) does occur before adulthood: a child with asthma, allergic rhinitis, or rhinosinusitis with polyps may respond to aspirin or another strong COX-1 inhibitor with rhinorrhea, sneezing, nasal blockage, ocular symptoms, cough, dyspnea, or marked bronchospasm, even when the complete Samter triad has not yet emerged.[5,6]

In addition, the accuracy of clinical history in suggesting NSAID hypersensitivity is relatively low in children. Only a third to a half of children labeled on clinical grounds are confirmed when challenged, so drug provocation testing (DPT) remains the reference method.[7-9] An incorrect label narrows the choice of analgesics and places an unnecessary burden on the family.

Data on NSAID hypersensitivity in children remain limited, and most published studies have primarily focused on the diagnostic accuracy of clinical history and on confirmation of suspected reactions by drug provocation testing. In contrast, information regarding long-term clinical outcomes and the distribution of respiratory phenotypes in the pediatric population is scarce. Therefore, the present study aimed to characterize the clinical features of children diagnosed with NSAID hypersensitivity at a single tertiary pediatric allergy center, determine the distribution of phenotypes according to the 2025 WAO classification, and assess the frequency of concomitant respiratory disease.

METHODS

Study Design and Patients

Children aged 1-18 years who were followed for NSAID hypersensitivity at the Pediatric Immunology and Allergy Clinic of University of Health Sciences Türkiye, İzmir Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital, Türkiye, between January 2015 and December 2020, were reviewed in this single-center, retrospective cross-sectional study. Eligibility required at least one reaction to NSAID intake. Children with an equivocal history and those whose families declined DPT were not included. The local ethics committee of University of Health Sciences Türkiye, Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital approved the protocol (June 3, 2021; approval no: 2021/10-05).

Data Collection

Demographic, clinical, and laboratory data were retrieved from the hospital electronic medical records and outpatient charts. The collected variables included age, sex, age at diagnosis, duration of follow-up, interval between the index reaction and diagnosis, family history of atopy and drug allergy, personal history of atopic disease, concomitant diseases, culprit NSAIDs, clinical characteristics of the hypersensitivity reaction, potential cofactors (infection or food intake), hospitalization due to the reaction, and treatments administered.

Laboratory evaluations included a complete blood count, an absolute eosinophil count (AEC), total serum IgE, allergen-specific IgE levels, and skin prick test results. Asthma control, assessed by the asthma control test (ACT), and current controller therapy were recorded for each child with asthma, and each patient was asked about nasal polyposis and prior sinus surgery.

Classification

Phenotypic classification was based primarily on a detailed clinical history and, when available, on DPT results, in accordance with the 2025 WAO classification; the defining features are summarized in Table 1. Because DPT was not performed in all patients, phenotypes assigned based clinical history were considered clinically probable rather than challenge-confirmed, unless hypersensitivity to the suspected culprit drug was demonstrated by a positive DPT. Urticaria, angioedema, or anaphylaxis after two or more chemically unrelated NSAIDs, without underlying chronic spontaneous urticaria, was recorded as NSAID-induced urticaria/angioedema/anaphylaxis (NIUAA); the same reactions confined to one drug or one chemical class, with other NSAIDs tolerated, were recorded as single NSAID-induced urticaria/angioedema/anaphylaxis (SNIUAA); symptoms arising only when a tolerated food was eaten together with an NSAID, were recorded as NSAID-exacerbated/induced food allergy (NEFA/NIFA); and delayed skin reactions after a single agent, were recorded as single NSAID-induced delayed hypersensitivity reactions (SNIDHR). Bronchospasm, cough, dyspnea, or a clear worsening of established airway symptoms after aspirin or a strong COX-1 inhibitor, once alternative causes had been excluded, define N-ERD, while exacerbation of pre-existing chronic spontaneous urticaria defines NSAID-exacerbated cutaneous disease (NECD).

NSAIDs were further categorized according to their chemical structure as propionic acid derivatives, salicylic acid derivatives, acetic acid derivatives, anthranilic acid derivatives (fenamates), pyrazolones, enolic acid derivatives (oxicams), para-aminophenol derivatives (paracetamol), sulfonanilide derivatives, and coxibs.

Drug Provocation Testing

Families were informed about DPT, and their consent was documented prior to the procedure. Challenges followed an open, incremental-dose protocol in a day-care unit equipped for resuscitation, and a skin prick test with the suspected drug preceded the challenge when feasible. DPT served either to confirm reactivity to the suspected agent or to establish a tolerated alternative. Children remained under observation for at least two hours after receiving the full cumulative dose. The test was stopped if any objective sign—urticaria, angioedema, bronchospasm, or hypotension—occurred; it was then recorded as positive and treated accordingly. Successive challenges in the same child were separated by at least one week. No patient underwent drug desensitization.

Table 1. WAO 2025 phenotypes of NSAID hypersensitivity[3]

Phenotype	Clinical manifestations	Timing	Underlying condition	Pathogenic mechanism	Cross-reactivity
NSAID-exacerbated respiratory disease (N-ERD)	Nasal congestion, rhinorrhea, sneezing, nasal itch, ocular symptoms, bronchospasm	Immediate (≤6 h)	Asthma, chronic rhinosinusitis and/or nasal polyps	COX-1 inhibition	Yes
NSAID-exacerbated cutaneous disease (NECD)	Urticaria, angioedema	Immediate (≤6 h)	Chronic spontaneous or recurrent urticaria	COX-1 inhibition	Yes
NSAID-induced urticaria/ angioedema/ anaphylaxis (NIUAA)	Urticaria, angioedema, anaphylaxis	Immediate (≤6 h)	None	COX-1 inhibition	Yes
Single-NSAID-induced urticaria/ angioedema/ anaphylaxis (SNIUAA)	Urticaria, angioedema, anaphylaxis	Immediate (≤6 h)	None	IgE-mediated	No
NSAID-exacerbated/induced food allergy (NEFA/NIFA)	Urticaria, angioedema, anaphylaxis	Immediate (≤6 h)	Food hypersensitivity or sensitization	NSAID as aggravating factor (NEFA) or cofactor (NIFA)	–
Single-NSAID-induced delayed hypersensitivity reactions (SNIDHR)	Contact dermatitis, FDE, MDE, SDRIFE, AGEP, DRESS, SJS, TEN, vasculitis	Delayed (>6 h)	None	T-cell mediated	No

NSAID: Non-steroidal anti-inflammatory drug, COX: Cyclooxygenase, FDE: Fixed drug eruption, MDE: Morbilliform drug eruption, SDRIFE: Symmetrical drug-related intertriginous and flexural exanthema, AGEP: Acute generalized exanthematous pustulosis, DRESS: Drug reaction with eosinophilia and systemic symptoms, SJS: Stevens-Johnson syndrome, TEN: Toxic epidermal necrolysis, WAO: World Allergy Organization

Follow-up and Clinical Outcomes

Parents or guardians were contacted by a structured telephone interview at the end of the study. Continued uneventful use of a previously documented alternative analgesic was recorded as safe alternative use, and uneventful re-exposure to the culprit drug was recorded as reported tolerance. Because these data were obtained by parental report, the dose administered, the number of subsequent exposures, and whether a full therapeutic dose was reached could not be verified; no confirmatory challenge was performed.

Statistical Analysis

SPSS 26.0 (IBM Corp., Armonk, NY, USA) was used for analysis. Continuous data are presented as median with range or interquartile range, and categorical data as counts and percentages. Groups were compared using the Mann-Whitney U test, the chi-square test, or the Fisher exact test, as required by the data, with significance set at a two-sided p below 0.05.

RESULTS

Demographic and Clinical Characteristics

Forty-three children were studied; 24 (55.8%) were boys. Median age at diagnosis was 7.8 years (range 1.3-17.5), median follow-up was 39 months (range 2-85), and median delay between the index reaction and diagnosis was 12 months (range 1-156). Diagnosis occurred between ages 2 and 5 years in 14 children (32.6%), between 6 and 11 years in 16 children (37.2%), and at 12 years or older in 12 children (27.9%); one child was diagnosed before the age of 2 years.

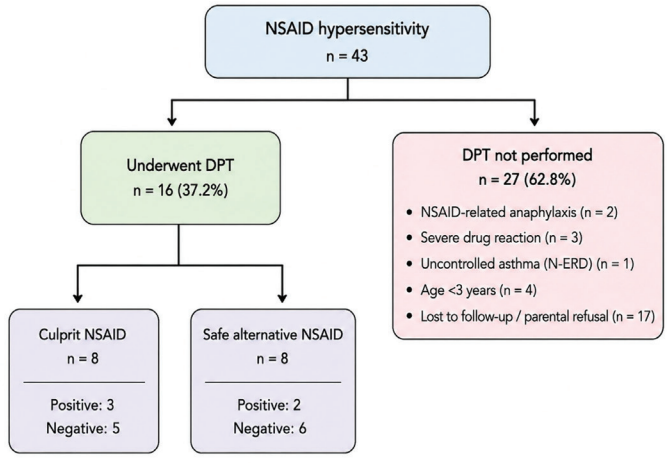


Figure 1. Flow diagram of drug provocation testing (DPT), showing eligibility, reasons for not testing, and the outcome of culprit and alternative NSAID challenges in 43 children with NSAID hypersensitivity.

NSAID: Non-steroidal anti-inflammatory drug

Twenty-one children (48.8%) had an atopic disorder; the most common were allergic rhinitis or rhinoconjunctivitis (n=10, 23.3%), followed by asthma (n=7), atopic dermatitis (n=2), and food allergy (n=2). Atopy in the family was recorded for 24 children (55.8%), and drug or NSAID hypersensitivity in a relative was recorded for 9 (20.9%). Infection accompanied the index reaction in six children (14.0%), and packaged food intake accompanied the index reaction in one child. Uneventful earlier use of the culprit

drug was reported in 38 of 41 children (92.7%), and 28 children (68.3%) had reacted on more than one occasion. Table 2 presents the full profile.

Reaction Characteristics and Phenotype Distribution

Urticaria, angioedema, or both accounted for most presentations (36 patients, 83.7%). Two children (4.7%) had anaphylaxis with hypotension, generalized urticaria, and vomiting. Three (7.0%) patients had severe delayed skin reactions, with one case each of DRESS, SJS, and AGEP. In two children (4.7%), asthma worsened after NSAID intake. Four (9.8%) were admitted because of the index reaction; no fatalities occurred.

Under the 2025 WAO scheme, SNIUAA was the most common phenotype (20 patients, 46.5%), followed by NIUAA in 18 patients (41.9%), SNIDHR in 3 patients (7.0%), and probable N-ERD in 2 patients (4.7%) (Table 3). Chronic spontaneous urticaria was absent from the cohort; therefore, no child met the criteria for NECD, and no NEFA/NIFA cases were observed. The phenotype groups did not differ in age at diagnosis, sex, atopy, follow-up length, AEC, or total IgE ($p>0.05$ throughout). Reactions to chemically unrelated NSAIDs were, predictably, far more common in NIUAA than in SNIUAA (94.4% vs. 5.0%, $p<0.001$).

Culprit Drugs

A specific agent was identified in 41 children (95.3%); the record read only "NSAID" for the other two. Ibuprofen was the most frequent, with 28 cases (65.1%), followed by paracetamol (n=18, 41.9%), metamizole (n=9, 20.9%), dexketoprofen (n=4, 9.3%), and acetylsalicylic acid (n=3, 7.0%). Culprit medications originated from a single chemical class in 22 children (53.7%) and from several classes in 19 children (46.3%). The most common pairing was paracetamol with ibuprofen, reported by 10 children.

Laboratory Findings

Median AEC was 230 cells/mm^3 (range 10-817; n=27), and median total IgE was 160 kU/L (range 3.9-456; n=21). Skin prick testing was performed on 23 children, of whom 11 (47.8%) reacted to at least one aeroallergen or food, with molds, house dust mite, and pollens accounting for most sensitizations. Prick testing with the culprit NSAID was negative throughout.

Drug Provocation Testing

Twenty-seven children were not challenged for the following reasons: previous NSAID anaphylaxis (n=2), severe drug reactions (n=3), poorly controlled asthma with N-ERD (n=1), age below 3 years (n=4), and loss to follow-up or withheld parental consent (n=17) (Figure 1).

Drug provocation tests were performed on 16 patients. Five of the eight children challenged with the suspected culprit tolerated it. All five had reported reactions to two or more chemically unrelated NSAIDs and had been assigned to the NIUAA phenotype on the basis of history. In each case, the negative challenge allowed removal of the label for the tested agent alone. Because the remaining implicated NSAIDs were not challenged, the NIUAA phenotype was retained and avoidance of the untested drugs continued. No phenotype was revised after DPT. Among

Table 2. Demographic and clinical characteristics of the study population (n=43)

Variable	n (%) or median (range)
Sex, male	24 (55.8)
Age at diagnosis, years	7.8 (1.3-17.5)
Follow-up duration, months	39 (2-85)
Interval between index reaction and diagnosis, months	12 (1-156)
Personal history of atopy	21 (48.8)
Allergic rhinitis/rhinoconjunctivitis	10 (23.3)
Asthma (with or without allergic rhinitis)	7 (16.3)
Atopic dermatitis	2 (4.7)
Food allergy	2 (4.7)
Family history of atopy	24 (55.8)
Family history of drug allergy	9 (20.9)
Asthma	7 (16.3)
Well controlled/partly controlled/uncontrolled	3/2/2
Nasal polyposis	0 (0)
Hospitalization due to the index reaction	4/41(9.8)*

*Data were available for 41 patients

Table 3. Phenotype distribution and baseline characteristics according to the 2025 WAO classification

Characteristic	SNIUAA (n=20)	NIUAA (n=18)	SNIDHR (n=3)	Probable N-ERD (n=2)
Male sex, n (%)	10 (50.0)	10 (55.6)	2 (66.7)	2 (100)
Age at diagnosis, years, median	7.9	7.8	6.3	13.5
Follow-up duration, months, median	40.5	27.5	76.0	39.0
Personal history of atopy, n (%)	13 (65.0)	6 (33.3)	0 (0)	2 (100)
Asthma, n (%)	3 (15.0)	2 (11.1)	0 (0)	2 (100)

SNIUAA: Single-NSAID-induced urticaria/angioedema/anaphylaxis, NIUAA: NSAID-induced urticaria/angioedema/anaphylaxis, SNIDHR: Single-NSAID-induced delayed hypersensitivity reactions, N-ERD: NSAID-exacerbated respiratory disease, WAO: World Allergy Organization

the eight patients challenged with a presumed safe alternative, six tolerated the tested drug, whereas two developed urticaria (Table 4). Thus, five patients were shown to tolerate a suspected culprit drug, whereas six were shown to tolerate an alternative NSAID; these outcomes were considered separately.

Long-term Follow-up and Clinical Outcomes

Six of the 28 children who had taken an NSAID since diagnosis (21.4%) had retaken the culprit drug without reporting a reaction and were recorded as having reported tolerance upon culprit-drug re-exposure. The dose taken, the number of subsequent exposures, and whether a full therapeutic dose was reached could not be verified, and no subjects underwent supervised re-challenge. Among the six children, four with SNIUAA and two with NIUAA, the median age at diagnosis was 6.3 years and the median follow-up was 51 months; neither age at diagnosis (p=0.575) nor follow-up duration (p=0.801) distinguished them from the rest of the cohort. The remaining 22 (78.6%), two of them with SNIDHR, had used a previously tolerated alternative without trouble. A further 15 children had avoided NSAIDs entirely and, therefore, could not be evaluated for subsequent tolerance to NSAIDs. This proportion therefore represents reported re-exposure outcomes in a selected subgroup and should not be interpreted as a remission or tolerance rate for the entire cohort.

Patients with NSAID-Exacerbated Respiratory Disease

Both children with N-ERD (4.7%) were boys, diagnosed at 11 and 16 years of age. Each was atopic and was taking an inhaled corticosteroid with a long-acting β_2-agonist; asthma was partly controlled in one and uncontrolled in the other. They had neither nasal polyps nor earlier sinus surgery.

The first patient developed a cough, dyspnea, and an asthma exacerbation approximately 3 hours after ibuprofen intake. At hospital presentation, cough and wheezing were documented, and spirometry showed an FEV_1 of 71% predicted. He reported one previous episode of similar respiratory symptoms following exposure to ibuprofen. His ACT score was 12. There were no clinical or laboratory findings suggestive of a concomitant infection, and no evidence of chronic rhinosinusitis. DPT was not performed, and the patient was subsequently lost to clinical follow-up. During the final telephone interview, he reported subsequent use of paracetamol without a reaction, and reported continued treatment with inhaled corticosteroid/long-acting β_2-agonist therapy.

The second patient developed a cough and dyspnea approximately 2 hours after exposure to both paracetamol and dexketoprofen. During evaluation in the emergency department, the patient's oxygen saturation was 92%, rhonchi were noted on auscultation, and the respiratory symptoms improved following salbutamol administration. Spirometry showed an FEV_1 of 75% predicted. He reported two previous episodes with similar respiratory symptoms following NSAID exposure. There was no clinical evidence of concomitant infection or chronic rhinosinusitis. A DPT with nimesulide was subsequently performed to identify a safe alternative and yielded a negative result.

Table 4. Drug provocation tests: tested drug, role, outcome and diagnosis after testing

#	Phenotype before DPT	Drug tested	Role	Result	Diagnosis after DPT
1	NIUAA	Paracetamol	Alternative	Positive (urticaria)	NIUAA
2	NIUAA	Metamizole	Culprit	Positive (urticaria)	NIUAA
3	NIUAA	Paracetamol	Culprit	Negative	NIUAA; paracetamol delabeled
4	SNIUAA	Paracetamol	Alternative	Negative	SNIUAA; safe alternative documented
5	NIUAA	Metamizole	Alternative	Positive (urticaria)	NIUAA
6	SNIUAA	Paracetamol	Alternative	Negative	SNIUAA; safe alternative documented
7	NIUAA	Paracetamol	Culprit	Positive (urticaria)	NIUAA
8	SNIUAA	Paracetamol	Alternative	Negative	SNIUAA; safe alternative documented
9	NIUAA	Paracetamol	Alternative	Negative	NIUAA; safe alternative documented
10	NIUAA	Ibuprofen	Alternative	Negative	NIUAA; safe alternative documented
11	NIUAA	Paracetamol	Culprit	Negative	NIUAA; paracetamol delabeled
12	NIUAA	Paracetamol	Culprit	Negative	NIUAA; paracetamol delabeled
13	NIUAA	Paracetamol	Culprit	Negative	NIUAA; paracetamol delabeled
14	NIUAA	Ibuprofen	Culprit	Positive (urticaria)	NIUAA
15	NIUAA	Paracetamol	Culprit	Negative	NIUAA; paracetamol delabeled
16	Probable N-ERD	Nimesulide	Alternative	Negative	Probable N-ERD; safe alternative documented

DPT: Drug provocation test, NSAID: Non-steroidal anti-inflammatory drug, SNIUAA: Single-NSAID-induced urticaria/angioedema/anaphylaxis, NIUAA: NSAID-induced urticaria/angioedema/anaphylaxis, N-ERD: NSAID-exacerbated respiratory disease,

DISCUSSION

We reviewed 43 children with NSAID hypersensitivity, assigned them to the 2025 WAO phenotypes, and followed their course. Cutaneous phenotypes predominated; N-ERD, although rare, was present in the cohort.

Skin involvement in 83.7% of our patients is consistent with earlier pediatric work. Podlecka et al.[10] found urticaria, angioedema, or both in 84.9% of confirmed cases, and Kidon et al.[5,11] described the same pattern as the usual presentation of cross-reactive hypersensitivity in younger children. Culprit distribution also agreed with published series: ibuprofen (65.1%) and paracetamol (41.9%) headed our list, which most likely reflects how often the two are given as antipyretics rather than any greater intrinsic risk.[7,12]

SNIUAA (46.5%) were more prevalent than NIUAA (41.9%) in this study. Arikoglu et al.[8] reported a similar ranking among 31 confirmed children—SNIUAA 11, NIUA 8, NECD 6, and N-ERD 1. Chronic spontaneous urticaria was absent from our cohort, so NECD did not appear; the disorder is less frequent in children than in adults, among whom Koh et al.[13] classified 11.6% of cross-reactive cases as NECD. NEFA/NIFA, which was added to the classification in 2025, was likewise unrepresented.

N-ERD cases accounted for 4.7% of our cohort; they were followed up as probable cases of N-ERD because DPT testing for the offending drug was not possible. Pediatric figures are sparse: roughly 0.9% of asthmatic children have confirmed NSAID hypersensitivity and about 0.3% have N-ERD.[6] The condition usually declares itself in the third or fourth decade, with reported onsets spanning puberty to the eighth decade.[14] Among adults with asthma it affects 5.5-12.4%, rising toward 20% where aspirin challenge is applied.[15,16] Diagnosis rests chiefly on a characteristic history and physical examination rather than on laboratory testing. Neither of our two patients had nasal polyps, in keeping with earlier behavior pediatric reports, which suggests early disease in which the Samter triad is not yet complete.[6,17] Asthma and upper airway symptoms, therefore, should be systematically assessed in any child suspected of NSAID hypersensitivity; prescribing NSAIDs to asthmatic children warrants caution.

Only 37.2% of our patients were challenged, which is the principal limitation of the study. Parental refusal and loss to follow-up limit confirmation in everyday practice, and other series from Türkiye report the same difficulty.[9] Nonetheless, the reference standard remains a challenge, since history alone misleads. Yilmaz et al.[7] confirmed just 14% of children reporting reactions to one NSAID and 44% of those reporting reactions to several. A recent prospective study of 336 children confirmed roughly one in three; febrile infection explained a large share of the remainder, and NEFA/NIFA accounted for 14%.[18] We identified no NEFA/NIFA cases. Infection coincided with the index reaction in 14.0% of our patients; therefore, some episodes may have been infection-related exanthems rather than true drug hypersensitivity. Labeling a child as NSAID-allergic without confirmation risks lifelong avoidance for no reason.[4,19]

Finding a well-tolerated analgesic is as important as identifying the cause. Paracetamol was tolerated in 9 of 11 challenges (81.8%), and nimesulide was tolerated in the one child tested. Yilmaz et al.[7] reported tolerance to paracetamol in 60% and to nimesulide in 88.8% of children, and adult data give tolerance rates of 90% for paracetamol 1,300 mg, 96% for meloxicam 15 mg, and 98% for celecoxib 200 mg in cross-reactive disease.[13] Agents with weak COX-1 activity and preferential or selective COX-2 inhibitors are thus reasonable choices for cross-reactive phenotypes, once the challenge has proved negative. Metamizole behaved differently; both challenges were positive; therefore, pyrazolones warrant caution as substitutes.

Tolerance upon re-exposure to the culprit drug was reported in 21.4% of children who had used an NSAID during follow-up. This finding should be interpreted differently from the 52.1% tolerance rate reported by Uluc et al.,[20] who established tolerance by supervised re-challenge after at least three years and reported a median time to tolerance of 6.16 years, with an initial reaction age of ≤11.75 years associated with tolerance development. In our study, tolerance was assessed using parental reports for children who were re-exposed, whereas those who continued to avoid NSAIDs could not be evaluated for subsequent tolerance. Therefore, our findings cannot be interpreted as an estimate of the spontaneous resolution rate of NSAID hypersensitivity. Supervised re-challenge, when clinically appropriate, may be valuable several years after diagnosis to identify children who have developed tolerance and to potentially avoid unnecessary long-term drug restriction.

Study Limitations

The main limitations of this study include its retrospective, single-center design, small sample size, and inability to perform DPT in all patients. An important consideration when interpreting our findings is that phenotype assignment was predominantly based on clinical history, because DPT could not be performed in all patients. Therefore, the reported phenotype distribution primarily reflects clinically assigned WAO phenotypes rather than challenge-confirmed NSAID hypersensitivity. In particular, the two respiratory cases were classified as probable N-ERD because culprit-drug DPTs were not performed, despite recurrent temporally related reactions and documented objective respiratory findings. Furthermore, long-term outcome data were obtained through telephone interviews and are therefore susceptible to recall bias. Despite these limitations, the relatively long median follow-up period of 39 months and the application of the 2025 WAO classification represent important strengths of the study.

CONCLUSION

Urticaria and angioedema characterize most pediatric NSAID hypersensitivity, with ibuprofen and paracetamol the most commonly reported culprits. Selective and cross-reactive patterns occurred at broadly similar frequencies, and although uncommon, airway involvement compatible with N-ERD was present in childhood and should be sought. Because phenotypes were assigned predominantly on clinical grounds, these proportions are provisional. Careful airway assessment, confirmation by challenge, wherever feasible, and documentation of a tolerated analgesic form the core of management. Prospective multicenter studies are needed to characterize the natural history and predictors of tolerance.

Ethics

Ethics Committee Approval: The local ethics committee of University of Health Sciences Türkiye, Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital approved the protocol (June 3, 2021; approval no: 2021/10-05).

Informed Consent: Written consent was obtained from parents or legal guardians before drug provocation testing. For the retrospective record review, the requirement for written consent was waived by the ethics committee, and verbal consent was obtained prior to the follow-up telephone interviews.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ö.S., Concept: Ö.S., F.Ç.Ç., Design: C.Ş.K., Ö.S., D.C., Data Collection or Processing: Ö.S., F.Ç.Ç., S.Ö.B., İ.T., M.Ş.K., İ.A.H., Analysis or Interpretation: Ö.S., Literature Search: Ö.S., Writing: Ö.S., F.Ç.Ç., C.Ş.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Blanca-Lopez N, Soriano V, Garcia-Martin E, Canto G, Blanca M. NSAID-induced reactions: classification, prevalence, impact, and management strategies. J Asthma Allergy. 2019;12:217-33.
2. Doña I, Pérez-Sánchez N, Eguiluz-Gracia I, et al. Progress in understanding hypersensitivity reactions to nonsteroidal anti-inflammatory drugs. Allergy. 2020;75:561-75.
3. Romano A, Valluzzi RL, Alvarez-Cuesta E, et al. Updating the classification and routine diagnosis of NSAID hypersensitivity reactions: aWAO statement. World Allergy Organ J. 2025;18:101086.
4. Steering Committee Authors; Berges-Gimeno MP, Alvarez-Cuesta E, et al. Drug hypersensitivity reactions in children in clinical practice: a WAO Statement. World Allergy Organ J. 2025;18:101087.
5. Kidon MI, Kang LW, Chin CW, Hoon LS, Hugo VB. Nonsteroidal anti-inflammatory drug hypersensitivity in preschool children. Allergy Asthma Clin Immunol. 2007;3:114-22.
6. Cavkaytar O, Arga M. NSAID hypersensitivity in the pediatric population: classification and diagnostic strategies. J Asthma Allergy. 2022;15:1383-99.
7. Yilmaz O, Ertoy Karagol IH, Bakirtas A, et al. Challenge-proven nonsteroidal anti-inflammatory drug hypersensitivity in children. Allergy. 2013;68:1555-61.
8. Arikoglu T, Aslan G, Yildirim DD, Batmaz SB, Kuyucu S. Discrepancies in the diagnosis and classification of nonsteroidal anti-inflammatory drug hypersensitivity reactions in children. Allergol Int. 2017;66:418-24.
9. Yilmaz Topal O, Kulhas Celik I, Turgay Yagmur I, et al. Results of NSAID provocation tests and difficulties in the classification of children with nonsteroidal anti-inflammatory drug hypersensitivity. Ann Allergy Asthma Immunol. 2020;125:202-7.
10. Podlecka D, Socha-Banasiak A, Jerzynska J, Nodzykowska J, Brzozowska A. Practical approach to hypersensitivity to nonsteroidal anti-ınflammatory drugs (NSAIDs) in children. Pharmaceuticals (Basel). 2023;16:1237.
11. Kidon MI, Kang LW, Chin CW, et al. Early presentation with angioedema and urticaria in cross-reactive hypersensitivity to nonsteroidal antiinflammatory drugs among young, Asian, atopic children. Pediatrics. 2005;116:e675-80.
12. Aytekin Güvenir F, Turgay Yağmur İ, Dibek Mısırlıoğlu E. Alternative drug safety in children with nonsteroidal anti-inflammatory drug hypersensitivity. Int Arch Allergy Immunol. 2024;185:921-7.
13. Koh YI, Yu JE, Sim DW. Cross-reactive NSAID hypersensitivity: clinical findings from aspirin provocation and alternative drug challenge testing. Clin Transl Sci. 2025;18:e70335.
14. Bensko JC, McGill A, Palumbo M, Cahill KN, Laidlaw TM. Pediatric-onset aspirin-exacerbated respiratory disease: clinical characteristics, prevalence, and response to dupilumab. J Allergy Clin Immunol Pract. 2022;10:2466-8.
15. Kowalski ML, Agache I, Bavbek S, et al. Diagnosis and management of NSAID-exacerbated respiratory disease (N-ERD)-a EAACI position paper. Allergy. 2019;74:28-39.
16. Haque R, White AA, Jackson DJ, Hopkins C. Clinical evaluation and diagnosis of aspirin-exacerbated respiratory disease. J Allergy Clin Immunol. 2021;148:283-91.
17. Kowalski ML, Asero R, Bavbek S, et al. Classification and practical approach to the diagnosis and management of hypersensitivity to nonsteroidal anti-inflammatory drugs. Allergy. 2013;68:1219-32.
18. Valluzzi RL, Urbani S, Sciannamea M, et al. Diagnosing nonsteroidal anti-inflammatory drug-hypersensitivity/allergy and nonsteroidal anti-inflammatory drug-exacerbated or induced food allergy phenotypes in children and adolescents. J Allergy Clin Immunol Pract. 2026:S2213-2198(26)00582-9.
19. Barbaud A, Garvey LH, Torres M, et al. EAACI/ENDA position paper on drug provocation testing. Allergy. 2024;79:565-79.
20. Uluc NN, Uluc NN, Iskender N, et al. Tolerance evaluation of nonsteroidal anti-inflammatory drug hypersensitivity in children: is age the crucial factor? Int Arch Allergy Immunol. 2025;186:1155-65.

DOI: 10.4274/forbes.galenos.2026.34713
Forbes J Med 2026;7:147-56

Original Article

The Evolution of Ambulatory Anesthesia Publications: A Bibliometric Analysis (1980-2025)

Selçuk KAYIR[1], Şeyda KAYHAN ÖMEROĞLU[2]

[1]Hitit University Faculty of Medicine, Department of Anesthesiology and Reanimation, Çorum, Türkiye
[2]Memorial Ataşehir Hospital, Medical Director, İstanbul, Türkiye

Cite as: Kayır S, Kayhan Ömeroğlu Ş. The evolution of ambulatory anesthesia publications: a bibliometric analysis (1980-2025). Forbes J Med. 2026;7:147-56

ABSTRACT

Objective: Ambulatory anesthesia has become a cornerstone of modern surgical care, yet its scientific literature has not previously been subjected to a comprehensive bibliometric evaluation. This study aimed to analyze global research output on ambulatory anesthesia between 1980 and 2025 and to identify publication trends, the most productive countries, institutions, authors, and journals, influential articles, and emerging research themes.

Methods: Documents published between 1980 and 2025 were retrieved from the Web of Science Core Collection using a structured Boolean search strategy. Only original articles and review articles were included. Bibliometric network visualizations were generated with VOSviewer, and statistical analyses, including cubic regression to future output, were performed.

Results: A total of 1,364 documents were retrieved; of these, 1,078 original articles and reviews were included. These publications received 32,104 citations in total, with an average of 29.78 citations per item and an h-index of 81. The United States was the most productive country, followed by England, Germany, Canada, and France; Türkiye ranked sixth. Anesthesia and Analgesia was the most productive journal, and White PF was the most prolific author. The most cited article addressed the etiology, treatment, and prevention of postoperative nausea and vomiting. Propofol, recovery, and short-acting volatile agents dominated the keyword landscape.

Conclusion: Research on ambulatory anesthesia has grown steadily over four decades and remains dominated by high-income countries. The findings underline the field's enduring focus on rapid recovery, antiemesis, and patient safety, providing a reference map for researchers, clinicians, and journals.

Keywords: Ambulatory anesthesia, day-case surgery, bibliometrics, scientometrics, publication trend analysis

Received: 03.07.2026
Accepted: 21.09.2026
Publication Date: 09.10.2026

Corresponding Author:
Selçuk KAYIR
Hitit University Faculty of Medicine, Department of Anesthesiology and Reanimation, Çorum, Türkiye
✉ drskayir@gmail.com
ORCID ID: 0000-0002-3176-7859

INTRODUCTION

Ambulatory anesthesia, also referred to as day-case or outpatient anesthesia, encompasses the anesthetic management of patients who are admitted, operated upon, and discharged on the same calendar day. Over the past four decades, driven by advances in short-acting anesthetic agents, minimally invasive surgical techniques, enhanced recovery pathways, and mounting economic pressure on hospital resources, an increasing proportion of surgical procedures worldwide has migrated to the ambulatory setting. The founding of the Society for Ambulatory Anesthesia in 1984 and the subsequent publication of dedicated consensus statements and day-case guidelines reflect the maturation of ambulatory anesthesia as a distinct subspecialty.

Bibliometrics is a scientific field that applies statistical and mathematical methods to the quantitative analysis of published literature and describes publication patterns within a defined area. First proposed by Pritchard in 1969 as a term to replace the ambiguous expression "statistical bibliography", bibliometrics enables objective measurement of the contributions of authors, institutions, and countries to a body of knowledge and the monitoring of the growth, development, and thematic evolution of scientific disciplines.[1] S cientometrics is defined as the quantitative study of the various aspects of science and technology.[2] Bibliometric indicators such as citation analysis, author and organization productivity, and keyword co-occurrence networks have become valuable tools for identifying influential works, leading contributors, and emerging research frontiers.

Although the volume of research on ambulatory anesthesia has increased substantially in recent decades, and bibliometric analyses have been published on several adjacent topics in

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

anesthesiology, including regional anesthesia and specific regional-block techniques, the literature currently lacks a holistic bibliometric evaluation devoted specifically to ambulatory anesthesia.[3,4] To the best of our knowledge, the present study is the first to comprehensively analyze the scientometric features of the ambulatory anesthesia literature. The aim of this study was to perform a bibliometric analysis of documents published in the field of ambulatory anesthesia between 1980 and 2025 and thereby demonstrate publication trends; the most productive countries, institutions, authors, and journals; the most influential articles; and the principal research themes and their evolution over time.

METHODS

Data Source and Search Strategy

All data analyzed in this study were retrieved from the Web of Science Core Collection (WoSCC; Clarivate Analytics) on a single day (2 July 2026) to avoid bias arising from continuous database updates. All citation metrics reflect this date. The following topic (TS) search strategy, which combines the principal synonyms for ambulatory anesthesia, was used: TS = ("ambulatory an*esthesia" OR "ambulatory an*esthetic*" OR "day-case an*esthesia" OR "day case an*esthesia" OR "day-surgery an*esthesia" OR "day surgery an*esthesia" OR "outpatient an*esthesia" OR "outpatient an*esthetic*" OR "office-based an*esthesia" OR "ambulatory surgery an*esthesia" OR "ambulatory surgical an*esthesia" OR "day-case surgery an*esthesia" OR "an*esthesia for ambulatory surg*" OR "an*esthesia for day-case*" OR "an*esthesia for day surg*" OR "an*esthesia for outpatient*").

The wildcard operator (an*esthesia) was used to capture both the American (anesthesia) and British (anaesthesia) spellings. All documents published between 1980 and 2025 were included. Papers published in 2026 were excluded to avoid bias introduced by an incomplete publication year. Publications from China, P.R.C., and the People's Republic of China were collected under a single "China" heading. Only original articles and review articles were included in the bibliometric analysis; the full distribution of document types is reported in the Results. Records dually indexed as "Article; Proceedings Paper" or "Article; Book Chapter" were retained, as these represent substantive articles and reviews rather than stand-alone conference abstracts.

Bibliometric Analysis and Visualization

Bibliometric network visualizations, including co-authorship networks among countries, co-citation networks, and keyword co-occurrence maps, were created using VOSviewer (version 1.6.21).[5] In the network maps, the more frequently occurring keywords, authors, and institutions formed larger nodes than those with lower frequency, and related or connected items were placed close to one another and colored according to cluster membership. Statistical analyses were performed using standard statistical software. A productivity score was calculated for each country using the previously reported formula (total number of publications/population × 1,000,000), an approach that has been applied in comparable holistic bibliometric investigations of other medical fields.[6-8] Cubic regression with curve fitting was used to model the cumulative growth of the literature because a third-order polynomial captured the non-linear S-shaped accumulation of publications more closely than linear or quadratic fits while avoiding overparameterization. The fitted model was then used to estimate the number of publications in the forthcoming year. A p value was considered statistically significant.

Ethics

This study was based exclusively on publicly available bibliographic data retrieved from the Web of Science Core Collection. The study did not involve the recruitment or interaction of human participants, access to individual-level patient or participant data, collection of human biological material, or the use of identifiable personal information. The bibliographic records analyzed in this study consisted of publication-level information, including article titles, authors, affiliations, publication years, journals, citation counts, and keywords. No interventions, clinical procedures, questionnaires, interviews, or other procedures involving human participants were performed. Accordingly, the study did not meet the definition of research involving human participants and did not require informed consent from individual participants. Furthermore, because the analysis involved only publicly available bibliographic records and did not involve human participants, human material, or animals, ethics committee approval was not required.

RESULTS

Total Number of Publications and Document Types

Our search of the ambulatory anesthesia literature retrieved 1,364 documents published between 1980 and 2025. Of these, 1,078 were original or review articles and were included in the bibliometric analysis. The vast majority of the analyzed documents were original articles (n=938), comprising 880 stand-

Table 1. Distribution of document types in the ambulatory anesthesia literature (1980-2025)

Document type	Number	%
Article	880	81.6
Review	138	12.8
Article; proceedings paper	49	4.5
Article; book chapter	9	0.8
Retracted publication	2	0.2
Total	1,078	100

Items indexed under more than one category (e.g., proceedings papers, book chapters) are counted once under their combined Web of Science label

alone articles, 49 items also indexed as proceedings papers, and 9 also indexed as book chapters; these were followed by review articles (n=138). Two of the included items (one article and one review) were subsequently retracted and are therefore flagged. Because certain records carry a combined Web of Science label (e.g., "Article; Proceedings Paper"), the subcategories in Table 1 are mutually exclusive and together sum to the analyzed total of 1,078.

By language, English was the dominant language of the ambulatory anesthesia literature, accounting for 991 documents (91.9%), followed by German (57, 5.3%), French (20, 1.9%), Turkish (4, 0.4%), Spanish (2), and Portuguese (2). The substantial German-language component reflects the historical contribution of the Central European anesthesia literature to the field.

Development of Publications and Citations

Figure 1 shows the distribution of publications and citations by year. An overall increasing trend in annual output was observed from 1980 onward, with the highest number of publications recorded in 2013 (n=40), followed by 2001 (n=39), 1997 and 2017 (n=38 each). The 1,078 analyzed publications were cited 32,104 times in total, with an average of 29.78 citations per item, and the h-index of the literature was calculated to be 81. The number of citations rose steeply in recent years, reaching its peak in 2025.

To model the growth of the field and estimate future output, a cubic curve was fitted to the cumulative number of publications (Figure 2). The model demonstrated an excellent fit (R^2=0.9984). Based on this model, the cumulative number of publications is estimated to reach 1,089 by the end of 2026-an increase of approximately 11 publications relative to the 1,078 publications recorded through 2025.

Productive Countries

Authors from more than 60 countries contributed to the ambulatory anesthesia literature. The United States was, by far, the most productive country with 442 publications, followed by England (101), Germany (85), Canada (71), and France (47). Notably, Türkiye ranked sixth, with 40 publications, ahead of China, Sweden, Italy, and Australia, indicating a considerable contribution from a middle-income country relative to more established research economies. The geographical distribution of the literature is presented in Table 2.

A productivity score, defined as the number of publications per million inhabitants, was calculated for the leading countries to adjust for differences in population size. When countries were ranked by productivity rather than by absolute output, the order changed considerably: Sweden (2.86) became the most productive country, followed by Canada (1.82), England (1.79), the United States (1.32) and Germany (1.01). This pattern, in which smaller, high-income countries achieve the highest per-capita output while the United States dominates in absolute numbers, is consistent with observations reported in bibliometric analyses of other medical fields.[9] Türkiye, despite its comparatively low gross domestic product per capita, achieved a productivity score (0.47) identical to that of Italy, again underscoring its notable contribution to the field. Productivity scores for the leading countries are presented in Table 3.

To examine international scientific collaboration, a country-level co-authorship network was constructed using VOSviewer (Figure 3). In this network, the size of each node reflects the number of publications, the thickness of the connecting lines represents the strength of collaboration, and the colors indicate clusters of countries that frequently published together. The United States formed by far the largest node and occupied a central position, connected to a wide range of partner countries. Canada and England emerged as two further major collaboration hubs, each anchoring its own cluster and linking, respectively, to partners such as Saudi Arabia and Thailand. Germany, France, Italy, Switzerland and China occupied intermediate positions, while Türkiye formed a distinct cluster of its own. The overall structure reveals a body of literature organized around a small number of predominantly

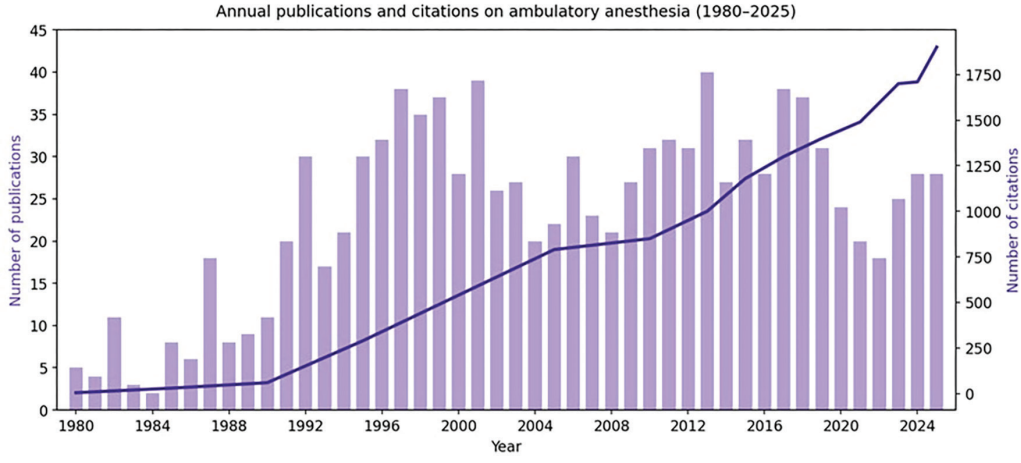


Figure 1. Number of publications (bars) and citations (line) on ambulatory anesthesia by year (1980-2025)

Table 2. The ten most productive countries in the ambulatory anesthesia literature	
Country	**Record count**
USA	442
England	101
Germany	85
Canada	71
France	47
Türkiye	40
China	30
Sweden	30
Italy	28
Australia	26

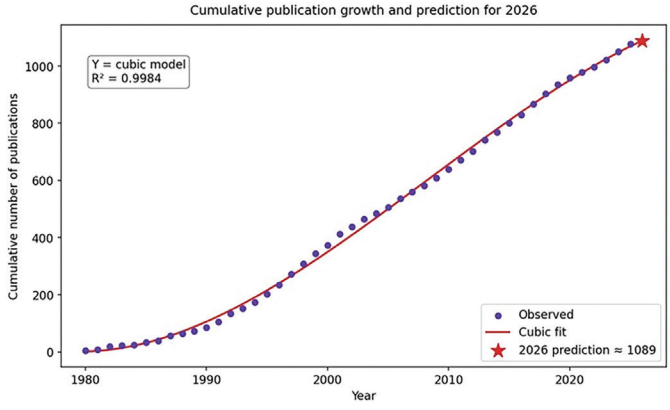


Figure 2. Cumulative number of ambulatory anesthesia publications by year, with cubic-curve fitting and prediction for 2026

high-income, highly collaborative countries, with the United States serving as the principal international collaboration hub.

Active Institutions

The most productive institutions in the ambulatory anesthesia literature were concentrated in North America. Harvard University and its affiliated medical institutions produced the highest number of publications, followed by the University of Texas System, the University of Toronto, the University of Texas Southwestern Medical Center, and the University of British Columbia. The prominence of these institutions is consistent with the leading positions held by authors affiliated with them, particularly White PF (University of Texas Southwestern) and Chung F (University of Toronto). The most productive organizations are presented in Table 4.

Table 3. Productivity scores (publications per one million inhabitants) of the leading countries	
Country	**Productivity score**
Sweden	2.86
Canada	1.82
England	1.79
USA	1.32
Germany	1.01
Australia	1.00
France	0.69
Türkiye	0.47
Italy	0.47
China	0.02

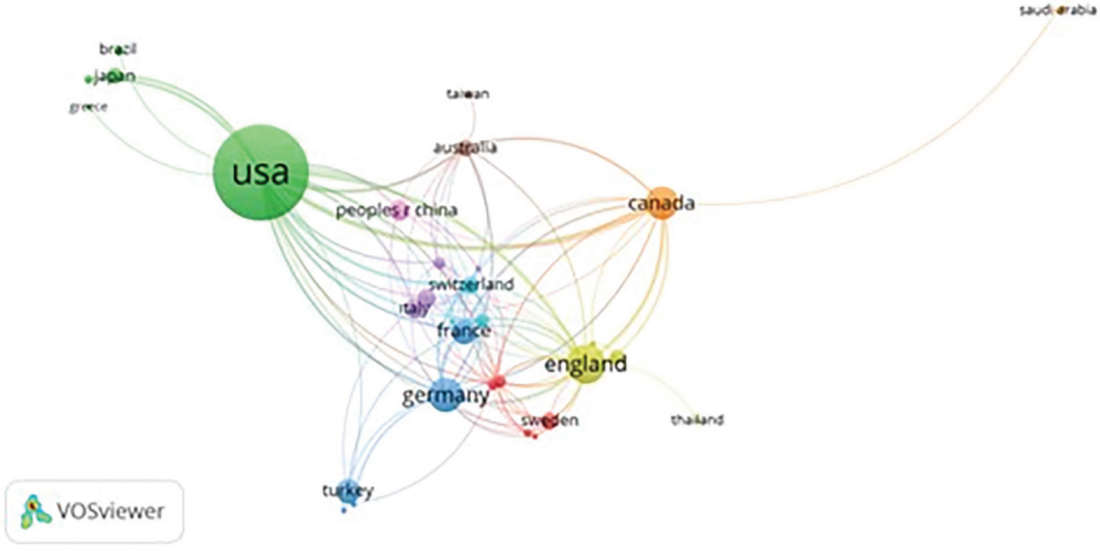


Figure 3. Co-authorship network of countries contributing to the ambulatory anesthesia literature. Node size reflects the number of publications; line thickness represents collaboration strength; colors indicate clusters

Active Authors

Based on the number of publications, White PF was by far the most prolific author in the ambulatory anesthesia literature (52 publications), reflecting his central role in the development of the field, particularly in research on recovery, fast-tracking and the pharmacology of short-acting anesthetics. He was followed by Chung F (24 publications), Smith I (20), Joshi GP (15), and Vaghadia H (13). Table 5 presents the leading authors.

Active Journals

A total of 1,078 articles were published across a large number of journals. Anesthesia and Analgesia was the leading journal with 109 publications, followed by Current Opinion in Anesthesiology (86), Acta Anaestesiologica Scandinavica (54), the Journal of Clinical Anesthesia (46), and the Journal of Oral and Maxillofacial Surgery (40). The prominence of the latter reflects the substantial body of ambulatory anesthesia research generated in office-based oral and maxillofacial surgery. The most productive journals are presented in Table 6.

Research Areas

As expected, anesthesiology was overwhelmingly the most frequent research area (733 publications), followed by general internal medicine (138), dentistry/oral surgery/medicine (80), surgery (71), and pediatrics (45). The strong representation of dentistry and oral surgery reflects the high volume of office-based and day-case procedures performed under sedation and general anesthesia in these disciplines. The leading research areas are presented in Table 7.

Most Cited Articles

The ten most frequently cited articles in the ambulatory anesthesia literature are presented in Table 8. The most cited document overall was the 1992 review by Watcha and White on the etiology, treatment, and prevention of postoperative nausea and vomiting (PONV) (1,241 citations); this topic recurs throughout the most influential works in the field, including successive consensus guidelines on its management.[10-12] When ranked by average citations per year, the 2022 American Society of Anesthesiologists practice guidelines for the management of the difficult airway achieved the highest impact (232.0 citations

Table 4. The most productive institutions in the ambulatory anesthesia literature

Institution	Record count
Harvard University	67
Harvard University Medical Affiliates	59
University of Texas System	57
University of Toronto	51
University of Texas Southwestern Medical Center	50
University of British Columbia	42
Harvard Medical School	40
University of Texas Dallas	35
University System of Ohio	35
University of California System	31

Table 5. The ten most productive authors in the ambulatory anesthesia literature

Author	Record count
White PF	52
Chung F	24
Smith I	20
Joshi GP	15
Vaghadia H	13
Shapiro FE	12
Urman RD	12
Gan TJ	11
Philip BK	11
Gupta A	9

Table 6. The ten most productive journals in the ambulatory anesthesia literature

Journal	Record count
Anesthesia and Analgesia	109
Current Opinion in Anaesthesiology	86
Acta Anaesthesiologica Scandinavica	54
Journal of Clinical Anesthesia	46
Journal of Oral and Maxillofacial Surgery	40
Anaesthesia	39
Anesthesiology	39
British Journal of Anaesthesia	37
Canadian Journal of Anaesthesia	35
European Journal of Anaesthesiology	33

Table 7. The ten most frequent research areas in the ambulatory anesthesia literature

Research area	Record count
Anesthesiology	733
General internal medicine	138
Dentistry oral surgery medicine	80
Surgery	71
Pediatrics	45
Pharmacology pharmacy	30
Neurosciences neurology	18
Orthopedics	16
Health care sciences services	13
Research experimental medicine	11

Table 8. The ten most cited articles in the ambulatory anesthesia literature

No	Article	First author	Year	TC	AC
1	Postoperative nausea and vomiting: its etiology, treatment, and prevention	Watcha MF	1992	1241	36.5
2	Consensus guidelines for the management of postoperative nausea and vomiting	Gan TJ	2014	1026	85.5
3	2022 American Society of Anesthesiologists practice guidelines for management of the difficult airway	Apfelbaum JL	2022	928	232.0
4	Fourth consensus guidelines for the management of postoperative nausea and vomiting	Gan TJ	2020	920	153.3
5	Society for Ambulatory Anesthesia guidelines for the management of postoperative nausea and vomiting	Gan TJ	2007	424	22.3
6	Can postoperative nausea and vomiting be predicted?	Sinclair DR	1999	418	15.5
7	Propofol: an update of its use in anesthesia and conscious sedation	Bryson HM	1995	293	9.4
8	A comparison of regional versus general anesthesia for ambulatory anesthesia: a meta-analysis of randomized controlled trials	Liu SS	2005	289	13.8
9	Titration of volatile anesthetics using bispectral index facilitates recovery after ambulatory anesthesia	Song DJ	1997	283	9.8
10	Comparison of recovery profile after ambulatory anesthesia with propofol, isoflurane, sevoflurane and desflurane: a systematic review	Gupta A	2004	268	12.2

TC: Total citations, AC: Average citations per year

per year), reflecting both the topicality of airway management and the rapid uptake of recent consensus guidelines.[13] PONV, consensus guidelines, recovery profiles of volatile agents, and regional versus general anesthesia emerged as the dominant themes among the most-cited articles.

Co-citation Analysis

A co-citation network was constructed to reveal the field's foundational knowledge base and internal structure (Figure 4). References that were frequently cited together were grouped into three principal clusters. A red cluster centered on the work of White PF, Gan TJ, Aldrete JA, and Gupta A, corresponding to the recovery, fast-tracking, and PONV axis that dominates modern literature; the widely cited modified Aldrete scoring reference and the Gan guideline references occupied central positions within this cluster. References concerning the recovery profiles of short-acting volatile and intravenous agents were grouped into a blue cluster (Song, Nathanson, Fredman, Tang). A green cluster comprised earlier, classic contributions to day-case anesthesia (Korttila, Doze, Tramer, Watcha, Mackenzie). The clear separation of these clusters illustrates the historical evolution of the field, from its foundational studies of anesthetic agents and recovery toward its current emphasis on standardized recovery assessment, antiemesis, and evidence-based guidelines.

Keyword Analysis and Research Themes

Keyword co-occurrence analysis was performed to reveal the principal research themes and their evolution (Figure 5). The most frequently used keywords were propofol (214 occurrences) and ambulatory anesthesia (214 occurrences), followed by surgery (186 occurrences), outpatient anesthesia (160 occurrences), anesthesia (158 occurrences,) and recovery (146 occurrences). The prominence of the short-acting intravenous agent propofol, together with the volatile agents sevoflurane (108), desflurane (101) and isoflurane (88), highlights the central and enduring concern of the field: the pharmacological achievement of rapid, high-quality recovery that permits same-day discharge. Postoperative nausea (78), sedation (64), pain (53), and regional anesthesia (45) constituted further important thematic clusters, while children (79) reflected the substantial pediatric ambulatory literature. The thirty most frequent keywords are presented in Table 9.

Table 9. The thirty most frequently used keywords in the ambulatory anesthesia literature

Keyword	Count	Keyword	Count
Propofol	214	Sedation	64
Ambulatory Anesthesia	214	Management	58
Surgery	186	Pain	53
Outpatient anesthesia	160	Complications	47
Anesthesia	158	Fentanyl	46
Recovery	146	Midazolam	45
General-anesthesia	123	Regional anesthesia	45
Sevoflurane	108	Maintenance	45
Ambulatory surgery	105	Bupivacaine	44
Desflurane	101	Alfentanil	43
Isoflurane	88	Remifentanil	41
Children	79	Safety	40
Postoperative nausea	78	Postoperative pain	37
Induction	68	Morbidity	37
Nitrous-oxide	68	Double-blind	37

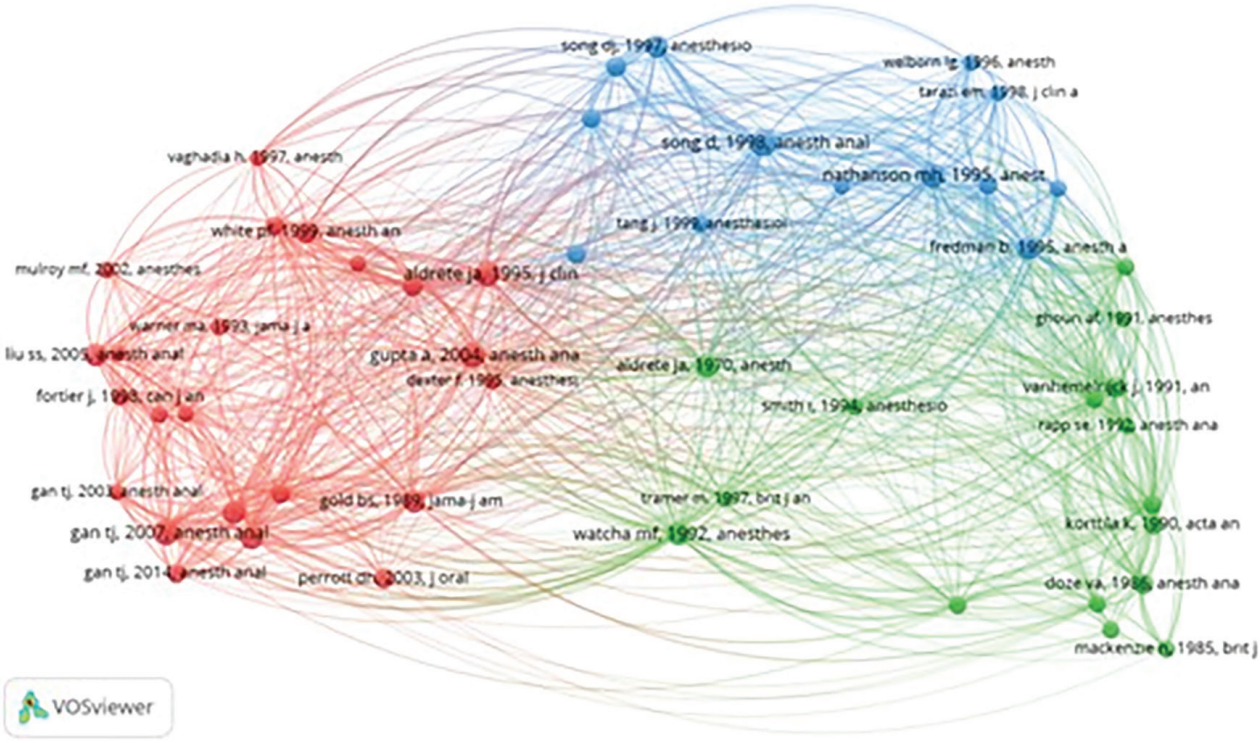


Figure 4. Co-citation network of cited references in the ambulatory anesthesia literature. Node size reflects co-citation frequency; colors indicate clusters of jointly cited references

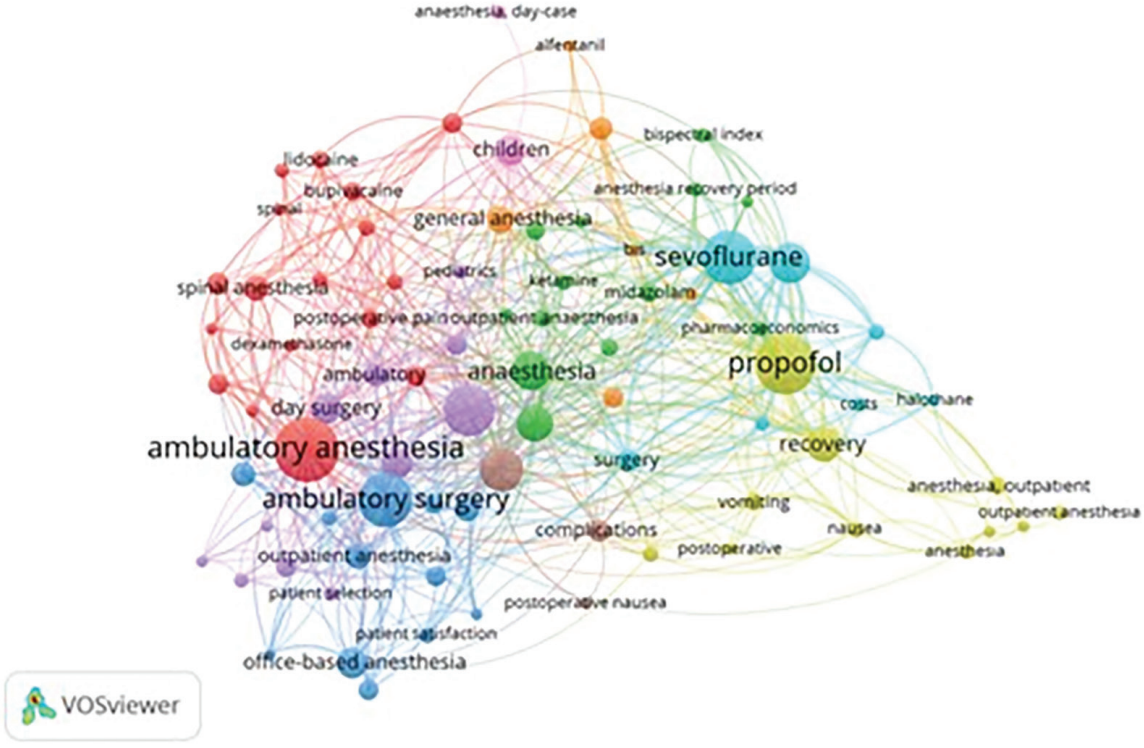


Figure 5. Keyword co-occurrence network of the ambulatory anesthesia literature. Node size reflects keyword frequency; colors indicate thematic clusters

DISCUSSION

Bibliometrics and scientometrics are established scientific methods that provide a quantitative, objective description of the publication and citation patterns within a defined field. Despite the growing clinical and economic importance of ambulatory surgery and the existence of bibliometric analyses on numerous related anesthesiology topics, the ambulatory anesthesia literature has not previously been evaluated holistically. To the best of our knowledge, the present study is the first comprehensive bibliometric analysis dedicated to this field.

The methodological approach adopted here-single-database extraction from the Web of Science Core Collection, calculation of population-adjusted productivity scores, network visualization with VOSviewer, and cubic-curve modeling of cumulative growth-follows a well-established tradition of holistic bibliometric evaluation that has been applied to a wide range of medical fields, including regional anesthesia, brain death, and legal medicine.[3,7,8] This consistency also extends beyond regional anesthesia to other anesthesiology subfields: recent global bibliometric analyses of opioid-free anesthesia and sufentanil research similarly identified the United States as the most productive and central contributor, PONV as a recurring thematic focus, and Anesthesia and Analgesia as among the most influential journals.[14,15] Positioning the present findings within this tradition allows comparison, in due course, of the ambulatory anesthesia literature with neighboring subspecialties. For example, a comparable bibliometric investigation of the regional anesthesia literature over a similar period also identified North American institutions and a small group of high-income countries as the principal contributors, suggesting that the concentration of research capacity observed here is characteristic of anesthesiology as a whole rather than specific to day-case practice.[3]

Our analysis of 1,078 original articles and reviews demonstrated that research output on ambulatory anesthesia grew steadily over the four decades studied, with an excellent fit to the cumulative growth curve and a projected continuation of this trend. The comparatively high average number of citations per item and an h-index of 81 indicate that, although this is a focused subspecialty, its literature exerts a substantial scientific impact that is concentrated in a relatively small number of highly cited guideline and review articles.

The United States dominated the ambulatory anesthesia literature in absolute terms, followed by England, Germany, Canada, and France. This distribution mirrors the general pattern observed across most medical bibliometric analyses, in which high-income, English-speaking countries lead the field.[9] However, when output was adjusted for population size through a productivity score, smaller high-income countries most notably Sweden, followed by Canada and England rose to the top of the ranking, while the United States remained highly productive but no longer occupied first place. Such a divergence between absolute and per-capita productivity has been repeatedly reported in the bibliometric literature, and it highlights the importance of complementing raw publication counts with population-adjusted measures.[6] A particularly noteworthy finding was Türkiye's sixth-place ranking in absolute output, ahead of several larger and wealthier research economies. This finding is consistent with the growing contribution of Turkish anesthesiology to the international literature and reflects the prominent Turkish presence documented in other medical bibliometric studies.[7]

The most productive institutions, Harvard University, the University of Texas System, the University of Toronto, and the University of British Columbia, reflected the leading positions of the individual authors affiliated with them. White PF was by far the most prolific author, an unsurprising finding given his foundational contributions to the pharmacology of short-acting anesthetics, recovery assessment, and the concept of fast-tracking after ambulatory surgery.[16,17] The most prolific authors after White PF, Chung F, Smith I, Joshi GP, and Vaghadia H have likewise sustained substantial programs of research in day-case anesthesia, ambulatory regional techniques, and recovery. Other influential contributors, including Gan TJ and Philip BK, are prominent for their work on PONV and the perioperative management of patients with obstructive sleep apnea, respectively.[18]

Anesthesia and Analgesia was the most productive journal, followed by Current Opinion in Anesthesiology and Acta Anaesthesiologica Scandinavica. The strong representation of the Journal of Oral and Maxillofacial Surgery, together with the prominence of dentistry and oral surgery as a leading research area, underscores an often-overlooked feature of the field: a large proportion of ambulatory anesthesia and sedation research is generated within office-based dental and oral-surgical practice. This observation may be of practical value to authors seeking suitable outlets for their work and to journals wishing to understand the disciplinary composition of their submissions.

The analysis of the most-cited articles, the co-citation network, and keyword co-occurrence produced a coherent picture of the intellectual structure of the field. The single most-cited document addressed PONV, and PONV-related consensus guidelines recurred among the most influential works, confirming that the prevention and treatment of PONV have been and remain one of the defining preoccupations of ambulatory anesthesia.[10-12] The co-citation analysis further clarified the historical architecture of the field, separating an earlier body of classic work on anesthetic agents and recovery from a more recent cluster organized around standardized recovery assessment, antiemesis, and evidence-based guidelines. In parallel, the dominance of propofol and the short-acting volatile agents sevoflurane, desflurane, and isoflurane among the most frequent keywords reflects the field's enduring central objective: to achieve anesthesia of a quality and rapidity of offset that permits safe same-day discharge.[19-21] More

recent high-impact publications, such as the 2022 difficult-airway guidelines and the 2019 day-case surgery guidelines, indicate that patient safety, standardization of care, and the expansion of day-case eligibility have become increasingly prominent themes.[13,22]

Several converging lines of inquiry emerge from these findings. The regional-versus-general debate, addressed by one of the most-cited meta-analyses in the field, continues to inform anesthetic choice for day-case surgery and points toward the expanding role of regional and plane-block techniques in the ambulatory setting.[23] Studies demonstrating that short-acting anesthetics can eliminate the need for intensive postoperative care and that fast-track eligibility can be predicted intraoperatively underpin the continued drive toward efficiency and same-day discharge.[21,24] At the same time, the recurrent appearance of research on PONV prediction and prolonged stays signals a persistent concern with identifying and managing patients least suited to an uncomplicated ambulatory pathway.[25-27] Future research is likely to continue with further refinement of opioid-sparing and multimodal analgesic strategies, expansion of regional techniques into the day-case setting, safe extension of ambulatory surgery to increasingly comorbid and elderly populations, and incorporation of enhanced-recovery and digital-health approaches. Researchers from middle- and lower-income countries should be encouraged and supported to contribute to this literature, so that the evidence base underpinning day-case anesthesia becomes more globally representative.

From a practical standpoint, this analysis offers several valuable insights. For researchers, it identifies the most influential works, the most active authors and institutions, and the journals most receptive to ambulatory anesthesia research, thereby assisting with literature review, collaboration, and manuscript placement. For journal editors, it clarifies the disciplinary composition of the field, most notably the substantial contribution of office-based dental and oral-surgical research, which may inform editorial strategy. For clinicians and trainees, the thematic map provides an accessible overview of the intellectual foundations of day-case anesthesia and the areas in which the evidence base is most and least developed.

Study Limitations

This study has certain limitations. First, we used only the Web of Science Core Collection and did not include other databases, such as Scopus, PubMed, or Google Scholar, which may index additional documents. We preferred the Web of Science because it is a highly reliable database that indexes journals with high impact factors and provides consistent, standardized bibliometric metadata. Second, as with all keyword-based bibliometric searches, the retrieved set is sensitive to the search strategy. Despite the use of a broad Boolean string combining the principal synonyms of ambulatory anesthesia, some relevant publications may not have been captured, and a small number of tangentially related documents may have been included. Third, citation counts are dynamic and reflect the state of the database on the date of extraction. These limitations are common to bibliometric research and are unlikely to have materially altered the principal findings. Finally, the forward projection should be interpreted with caution: a single-year extrapolation, derived from a cubic fit to the cumulative curve, is sensitive to the most recent data points and therefore carries considerable uncertainty, particularly because citations and indexing for the most recent years remain incomplete. Because cumulative publication counts naturally increase over time, a high R^2 should not be interpreted as evidence of predictive superiority; the projected value for 2026 should be regarded as an exploratory estimate rather than a precise forecast.

CONCLUSION

The scientific literature on ambulatory anesthesia has grown steadily over the past four decades and is projected to continue expanding. The field remains dominated by high-income countries led in absolute terms by the United States and, on a per-capita basis, by smaller nations such as Sweden, Canada and England while Türkiye stands out as a notable middle-income contributor. Research output has been concentrated in a small number of leading institutions, authors, and journals and has consistently centered on achieving rapid, high-quality recovery, preventing PONV, and maintaining patient safety. As the first comprehensive bibliometric analysis of this field, the present study provides a reference map of its intellectual structure and may assist researchers, clinicians, and journals in identifying influential works, active contributors, and promising directions for future investigation.

Ethics

Ethics Committee Approval: This study was based exclusively on publicly available bibliographic data retrieved from an online citation database and did not involve human participants, human material, or animals. Therefore, ethics committee approval was not required.

Informed Consent: Not applicable. This study used only publicly available, publication-level bibliographic data and did not involve human participants, participant interaction, individual-level personal data, or human biological material. Therefore, informed consent was not required.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.K., Concept: S.K., Ş.K.Ö., Design: S.K., Ş.K.Ö., Data Collection or Processing: S.K., Analysis or Interpretation: S.K., Ş.K.Ö., Literature Search: S.K., Writing: S.K., Ş.K.Ö.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Pritchard A. Statistical bibliography or bibliometrics? J Doc. 1969;25:348-9.
2. Hood WW, Wilson CS. The literature of bibliometrics, scientometrics, and informetrics. Scientometrics. 2001;52:291-314.
3. Kayir S, Kisa A. The evolution of the regional anesthesia: a holistic investigation of global outputs with bibliometric analysis between 1980-2019. Korean J Pain. 2021;34:82-93.
4. Coviello A, Iacovazzo C, Cirillo D, et al. The twenty-five most cited articles about adductor canal block: a bibliometric analysis from 1980 to 2022. J Pain Res. 2023;16:3353-65.
5. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics. 2010;84:523-38.
6. Şenel E, Demir E. Bibliometric analysis of apitherapy in complementary medicine literature between 1980 and 2016. Complement Ther Clin Pract. 2018;31:47-52.
7. Doğan G, Kayır S. Global scientific outputs of brain death publications and evaluation according to the religions of countries. J Relig Health. 2020;59:96-112.
8. Demir E, Yaşar E, Özkoçak V, Yıldırım E. The evolution of the field of legal medicine: a holistic investigation of global outputs with bibliometric analysis. J Forensic Leg Med. 2020;69:101885.
9. Şenel E. Evolution of homeopathy: a scientometric analysis of global homeopathy literature between 1975 and 2017. Complement Ther Clin Pract. 2019;34:165-73.
10. Watcha MF, White PF. Postoperative nausea and vomiting. Its etiology, treatment, and prevention. Anesthesiology. 1992;77:162-84.
11. Gan TJ, Diemunsch P, Habib AS, et al. Consensus guidelines for the management of postoperative nausea and vomiting. Anesth Analg. 2014;118:85-113.
12. Gan TJ, Belani KG, Bergese S, et al. Fourth consensus guidelines for the management of postoperative nausea and vomiting. Anesth Analg. 2020;131:411-48.
13. Apfelbaum JL, Hagberg CA, Connis RT, et al. 2022 American Society of Anesthesiologists practice guidelines for management of the difficult airway. Anesthesiology. 2022;136:31-81.
14. He J, Huang R, Liu Y, Chen Y, Zhong M. Global research frontiers and thematic trends in opioid-free anesthesia over the past 20 years: a bibliometric analysis. Front Pharmacol. 2025;16:1562765.
15. Huang D, Luo Z, Song X, Zou K. Global research on sufentanil use in anesthesiology from 2003 to 2023: a bibliometric analysis. Front Pharmacol. 2024;15:1412726.
16. White PF, Song D. New criteria for fast-tracking after outpatient anesthesia: a comparison with the modified Aldrete's scoring system. Anesth Analg. 1999;88:1069-72.
17. White PF, Eng M. Fast-track anesthetic techniques for ambulatory surgery. Curr Opin Anaesthesiol. 2007;20:545-57.
18. Joshi GP, Ankichetty SP, Gan TJ, Chung F. Society for Ambulatory Anesthesia consensus statement on preoperative selection of adult patients with obstructive sleep apnea scheduled for ambulatory surgery. Anesth Analg. 2012;115:1060-8.
19. Gupta A, Stierer T, Zuckerman R, Sakima N, Parker SD, Fleisher LA. Comparison of recovery profile after ambulatory anesthesia with propofol, isoflurane, sevoflurane and desflurane: a systematic review. Anesth Analg. 2004;98:632-41, table of contents.
20. Nathanson MH, Fredman B, Smith I, White PF. Sevoflurane versus desflurane for outpatient anesthesia: a comparison of maintenance and recovery profiles. Anesth Analg. 1995;81:1186-90.
21. Song D, van Vlymen J, White PF. Is the bispectral index useful in predicting fast-track eligibility after ambulatory anesthesia with propofol and desflurane? Anesth Analg. 1998;87:1245-8.
22. Bailey CR, Ahuja M, Bartholomew K, et al. Guidelines for day-case surgery 2019: Guidelines from the Association of Anaesthetists and the British Association of Day Surgery. Anaesthesia. 2019;74:778-92.
23. Liu SS, Strodtbeck WM, Richman JM, Wu CL. A comparison of regional versus general anesthesia for ambulatory anesthesia: a meta-analysis of randomized controlled trials. Anesth Analg. 2005;101:1634-42.
24. Apfelbaum JL, Walawander CA, Grasela TH, et al. Eliminating intensive postoperative care in same-day surgery patients using short-acting anesthetics. Anesthesiology. 2002;97:66-74.
25. Sinclair DR, Chung F, Mezei G. Can postoperative nausea and vomiting be predicted? Anesthesiology. 1999;91:109-18.
26. Apfel CC, Korttila K, Abdalla M, et al. A factorial trial of six interventions for the prevention of postoperative nausea and vomiting. N Engl J Med. 2004;350:2441-51.
27. Chung F, Mezei G. Factors contributing to a prolonged stay after ambulatory surgery. Anesth Analg. 1999;89:1352-9.

DOI: 10.4274/forbes.galenos.2026.57984
Forbes J Med 2026;7:157-64

Original Article

Preoperative Surgical Fear, Anxiety, Depression, and Pain in Orthopedic Patients: A Cross-Sectional Comparison of Trauma and Elective Surgery

Hamit Çağlayan KAHRAMAN[1], Burcu BAKAR KAHRAMAN[2], Ahmet Selim KARAMOLLA[1], Oğuzhan TANOĞLU[3], Barış YILMAZ[1]

[1]University of Health Sciences Türkiye, Fatih Sultan Mehmet Training and Research Hospital, Clinic of Orthopedics and Traumatology, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Sultan 2. Abdulhamid Han Training and Research Hospital, Clinic of Psychiatry, İstanbul, Türkiye
[3]İzmir Democracy University Faculty of Medicine, Buca Seyfi Demirsoy Training and Research Hospital, Department of Orthopedics and Traumatology, İzmir, Türkiye

Cite as: Kahraman HÇ, Bakar Kahraman B, Karamolla AS, Tanoğlu O, Yılmaz B. Preoperative surgical fear, anxiety, depression, and pain in orthopedic patients: a cross-sectional comparison of trauma and elective surgery. Forbes J Med. 2026;7:157-64

ABSTRACT

Objective: Orthopedic surgery patients may experience substantial physical and psychological burden, including pain, anxiety, depressive symptoms, and surgical fear, which may influence perioperative recovery. To compare preoperative surgical fear between trauma-related and elective orthopedic surgery patients and to examine associations of short- and long-term surgical fear with anxiety, depressive symptoms, and pain.

Methods: Consecutive patients aged ≥18 years undergoing trauma-related or elective orthopedic surgery at University of Health Sciences Türkiye, Fatih Sultan Mehmet Training and Research Hospital were enrolled. Participants completed the surgical fear questionnaire (SFQ), the hospital anxiety and depression scale (HADS), and the visual analog scale (VAS) one day before surgery.

Results: A total of 142 patients were analyzed (trauma, n=64; elective, n=78). Preoperative length of stay and VAS scores were higher in the trauma group. No statistically significant differences between groups were detected for short-term, long-term, or total SFQ scores. Short-term fear correlated with HADS anxiety (ρ=0.480, q<0.001), whereas long-term fear correlated with both HADS anxiety (ρ=0.476, q<0.001) and HADS depression (ρ=0.435, q<0.001). In multivariable analyses using HC3 robust standard errors, higher HADS anxiety and female sex were independently associated with greater short-term fear, whereas higher HADS anxiety and HADS depression were independently associated with greater long-term fear. Trauma/elective status was not independently associated with either of the fear dimensions.

Conclusion: Short- and long-term surgical fear were more strongly associated with psychological characteristics than with surgical indication. Anxiety was associated with both fear dimensions; female sex was associated with short-term fear, and depressive symptoms were associated with long-term fear. These findings support consideration of psychological assessment in preoperative orthopedic care.

Keywords: Trauma, arthroplasty, surgery, fear, anxiety, depression

Received: 13.04.2026
Accepted: 24.09.2026
Publication Date: 09.10.2026

Corresponding Author:
Hamit Çağlayan KAHRAMAN
University of Health Sciences Türkiye, Fatih Sultan Mehmet Training and Research Hospital, Clinic of Orthopedics and Traumatology, İstanbul, Türkiye
drhcaglayan@gmail.com
ORCID ID: 0000-0001-6150-9300

INTRODUCTION

Orthopedics and traumatology encompass a wide range of surgical procedures, and patients may experience substantial physiological and psychological stress throughout treatment and recovery. Factors influencing outcomes in orthopedic patients are not limited to implant selection, fracture characteristics, or surgical technique. Pain, adherence to rehabilitation, patient satisfaction, and functional recovery may also be influenced by psychological factors. Orthopedic injuries can impose a considerable physical and psychological burden, contributing to pain, functional limitations, and impaired quality of life.[1,2] Understanding the psychological burden associated with orthopedic trauma is therefore important for improving perioperative care and rehabilitation.[3]

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

Surgical fear includes short- and long-term concerns related to surgery, anesthesia, pain, postoperative adverse effects, complications, loss of independence, recovery, and unfavorable outcomes. The surgical fear questionnaire (SFQ) was developed to assess these dimensions systematically and has subsequently been adapted and validated in Turkish surgical populations.[4,5] Identifying factors associated with greater preoperative surgical fear may help clinicians recognize patients who could benefit from additional psychological assessment or perioperative support.

Illness and surgical treatment may also contribute to anxiety and depressive symptoms. Pronounced preoperative anxiety has been associated with adverse perioperative experiences and poorer recovery-related outcomes.[6,7] Pain may further interact with emotional distress and surgical fear, particularly in orthopedic patients in whom pain frequently represents a major component of the presenting condition. Previous studies have evaluated surgical fear in orthopedic and other surgical populations,[8-10] but direct comparisons between trauma-related and elective orthopedic surgery remain limited.

We hypothesized that patients undergoing trauma-related orthopedic surgery would have higher short-term surgical fear than those undergoing elective orthopedic surgery because of the acute onset, uncertainty, and sudden functional impairment associated with traumatic injury. The primary objective of this study was to compare short-term preoperative surgical fear between patients undergoing trauma-related orthopedic surgery and those undergoing elective orthopedic surgery. The secondary objectives were to compare long-term surgical fear between the groups and to examine its associations with anxiety, depressive symptoms, preoperative pain, and selected demographic and clinical characteristics.

METHODS

Ethical approval was obtained from the Ethical approval was obtained from the University of Health Sciences Türkiye, Koşuyolu Yüksek İhtisas Training and Research Hospital Ethics Committee (decision no: 2026/04/1386, date: 24.02.2026), as our hospital did not have a local ethics committee. Participant recruitment began after ethical approval had been obtained. The study was conducted in accordance with the Declaration of Helsinki and written informed consent was obtained from all participants prior to participation in the study. All study data were anonymized and used solely for scientific purposes.

The inclusion criteria were age ≥18 years; sufficient cognitive ability to comprehend and respond to the study instruments; ability to read and understand Turkish; willingness to participate voluntarily; and an appropriate preoperative clinical condition to complete the data collection form, SFQ, hospital anxiety and depression scale (HADS), and visual analog scale (VAS).

The exclusion criteria were emergency surgery, including cases carrying the American Society of Anesthesiologists (ASA) physical status "E" emergency modifier; age <18 years; severe cognitive impairment or a neurological or psychiatric condition interfering with effective communication; inability to read or understand Turkish; refusal to participate; incomplete study scales; inability to undergo standard preoperative assessment because of clinical instability; and missing study data.

This single-center, cross-sectional, observational study included 142 consecutive adult patients undergoing trauma-related or elective orthopedic surgery at the Orthopedics and Traumatology Clinic between February 26, 2026, and April 30, 2026. All assessments were performed at a single time point, one day before surgery.

A standardized data collection form was used to record demographic and clinical characteristics, including age, sex, educational level, marital status, number of children, place of residence, occupation, income level, caregiver or companion status, previous hospitalization, previous surgery, chronic diseases, ASA physical status classification, length of preoperative hospital stay, and preoperative pain intensity. Pain intensity was assessed using a 0-10 VAS.

The SFQ, developed by Theunissen et al.[4] and adapted into Turkish by Bağdigen and Karaman Özlü,[5] is an eight-item instrument assessing short- and long-term surgical fear. Each item is scored from 0 to 10, resulting in a total score ranging from 0 to 80, with higher scores indicating greater surgical fear. Four items assess short-term fear related to the immediate surgical and early perioperative period, while four assess long-term concerns related to recovery, surgical outcomes, and possible long-term consequences.

The HADS was developed for the assessment of anxiety and depressive symptoms in medical populations while minimizing the contribution of somatic symptoms. Its validity and reliability have been established in the Turkish population.[11] The HADS consists of 14 self-report items divided into two seven-item subscales: HADS-anxiety (HADS-A) and HADS-depression (HADS-D). Each subscale ranges from 0 to 21, with higher scores indicating greater symptom severity.

SFQ, HADS, and VAS assessments were completed during hospitalization in the patients' rooms under standardized conditions. In the elective surgery group, ASA classification was established before admission. In patients undergoing trauma-related surgery, ASA classification was determined during hospitalization following preoperative anesthetic evaluation. Preoperative length of hospital stay was defined as the interval from hospital admission to the day of surgery.

An a priori sample size analysis was performed using G*Power version 3.1. The calculation was based on the primary comparison of surgical fear between the trauma and elective groups. The sample size calculation was based on the between-group comparison of the primary short-term SFQ score. For a two-

sided comparison between two independent groups, assuming a medium standardized effect size (Cohen's d=0.50), an alpha level of 0.05, and statistical power of 80%, the required total sample size was 128 participants (64 per group). A larger sample was targeted to account for possible exclusions or incomplete data.

Statistical Analysis

Statistical analyses were performed using (version 24.0; IBM Corp., Armonk, NY, USA), with additional Python-based routines for robust inference, bootstrap confidence intervals (CIs), and false discovery rate adjustment. Continuous variables were assessed for normality and summarized as mean ± standard deviation or median (IQR), as appropriate; categorical variables were reported as n (%).

Between-group comparisons were performed using the Mann-Whitney U test for continuous variables and the chi-square test or, when required, exact tests for categorical variables. Rank-biserial correlations with bootstrap 95% CIs were reported as effect sizes. Associations among age, HADS-A, HADS-D, SFQ scores, and VAS were evaluated using Spearman correlation coefficients, with Benjamini-Hochberg correction applied to the 15 exploratory comparisons.

Separate multivariable linear regression models were constructed for short- and long-term surgical fear using age, sex, marital status, previous surgery, preoperative length of stay, trauma/elective status, HADS-D, HADS-A, and VAS as covariates. For regression analyses, sex was coded as female =0 and male =1; marital status as single =0 and married =1; surgical group as elective =0 and trauma =1; and previous surgery as none =0 and one or more previous operations =1. Regression assumptions were assessed using residual diagnostics, VIF and tolerance, the Ramsey RESET test, the Breusch-Pagan test, and Cook's distance. Because heteroscedasticity was detected in the long-term model, HC3 robust standard errors were used for both models. Sensitivity analyses included ASA classification and caregiver status. All tests were two-sided; $p<0.05$ was considered statistically significant, while exploratory correlations were interpreted based on adjusted $q<0.05$.

RESULTS

A total of 142 patients met the eligibility criteria and were included in the analysis, comprising 64 undergoing trauma-related orthopedic surgery and 78 undergoing elective orthopedic surgery.

Demographic and clinical characteristics are presented in Table 1. Significant between-group differences were observed in marital status, number of children, caregiver status, history of hospitalization, history of surgery, and ASA classification. In contrast, sex, education, place of residence, occupation, income level, and chronic disease status did not differ significantly between the groups.

Table 1. Comparison of demographic and clinical characteristics between the trauma and elective surgery groups

Variable	Trauma, n (%)	Elective, n (%)	p
Sex			0.145
Female	35 (54.7)	52 (66.7)	
Male	29 (45.3)	26 (33.3)	
Education level			0.222
Primary school	24 (37.5)	38 (48.7)	
Middle school	10 (15.6)	12 (15.4)	
High school	22 (34.4)	15 (19.2)	
Undergraduate degree	8 (12.5)	11 (14.1)	
Master's degree	0	2 (2.6)	
Marital status			<0.001
Single	34 (53.1)	15 (19.2)	
Married	30 (46.9)	63 (80.8)	
Number of children			0.020
None	26 (40.6)	16 (20.5)	
One child	10 (15.6)	13 (16.7)	
Two children	17 (26.6)	20 (25.6)	
≥3 children	11 (17.2)	29 (37.2)	
Place of residence			0.112
City	62 (96.9)	69 (88.5)	
District/town/village	2 (3.1)	9 (11.6)	
Occupation			0.050
Housewife	19 (29.7)	40 (51.3)	
Worker	6 (9.4)	7 (9.0)	
Self-employed	11 (17.2)	9 (11.5)	
Retired	10 (15.6)	13 (16.7)	
Unemployed	8 (12.5)	2 (2.6)	
Other	10 (15.7)	7 (8.9)	
Income level			0.671
Low	5 (7.8)	8 (10.3)	
Middle	57 (89.1)	69 (88.5)	
High	2 (3.1)	1 (1.3)	
Caregiver/companion			0.001
None	9 (14.1)	0	
Spouse	18 (28.1)	27 (34.6)	
Child	11 (17.2)	29 (37.2)	
Parent(s)	14 (21.9)	10 (12.8)	
Other	12 (18.8)	12 (15.4)	
History of previous hospitalization			<0.001
None	26 (40.6)	16 (20.5)	
Once	6 (9.4)	28 (35.9)	
Twice	9 (14.1)	11 (14.1)	
≥3 times	23 (35.9)	23 (29.5)	

Table 1. Continued			
Variable	**Trauma, n (%)**	**Elective, n (%)**	**p**
History of previous surgery			0.039
None	27 (42.2)	20 (25.6)	
Once	14 (21.9)	31 (39.7)	
Twice	23 (35.9)	27 (34.6)	
History of chronic disease			0.081
No	34 (53.1)	30 (38.5)	
Yes	30 (46.9)	48 (61.5)	
ASA classification			<0.001
ASA I	9 (14.1)	9 (11.5)	
ASA II	25 (39.1)	28 (35.9)	
ASA III	8 (12.5)	36 (46.2)	
ASA IV	22 (34.4)	5 (6.4)	

Data are presented as n (%). Pearson's chi-square test was used when assumptions were satisfied. Fisher's exact test or Fisher-Freeman-Halton exact testing was used for sparse tables. Exact/Monte Carlo testing was used for education level, place of residence, occupation, income level, and caregiver/companion status. A two-sided p<0.05 was considered statistically significant ASA: American Society of Anesthesiologists

Preoperative length of hospital stay was significantly longer in the trauma group than in the elective group [3 (2-3) vs. 1 (1-1) days, p<0.001], with a large rank-biserial effect size (r_rb=0.838, 95% CI 0.743 to 0.922). Preoperative VAS scores were also higher in the trauma group [7 (6-9) vs. 6 (4-7), p=0.001; r_rb=0.322, 95% CI 0.147 to 0.498].

No statistically significant between-group difference was detected in short-term surgical fear [23 (12-24) vs. 17 (10-25), p=0.175; r_rb=0.132, 95% CI -0.051 to 0.324], long-term surgical fear [20.5 (10.5-27) vs. 16.5 (11.5-24), p=0.112; r_rb=0.155, 95% CI -0.039 to 0.346], or total surgical fear [44 (28.25-50) vs. 34 (18-47), p=0.095; r_rb=0.163, 95% CI -0.030 to 0.354]. HADS-A, HADS-D, and total HADS scores also did not differ significantly between the groups (Table 2).

In exploratory Spearman correlation analyses with Benjamini–Hochberg adjustment, HADS-D was positively associated with short-term surgical fear (ρ=0.368, q<0.001) and long-term surgical fear (ρ=0.435, q<0.001). HADS-A was positively associated with both short-term fear (ρ=0.480, q<0.001) and long-term fear (ρ=0.476, q<0.001). Short- and long-term surgical fears were strongly correlated (ρ=0.739, q<0.001).

The VAS score showed weak positive correlations with HADS-D (ρ=0.190, q=0.030), HADS-A (ρ=0.322, q<0.001), short-term fear (ρ=0.198, q=0.025), and long-term fear (ρ=0.205, q=0.022). Age was weakly associated with short-term fear (ρ=0.206, q=0.022) and with VAS score (ρ=0.226, q=0.013), whereas its association with HADS-A did not remain statistically significant after FDR adjustment (ρ=0.160, q=0.067) (Table 3).

Regression diagnostics showed no evidence of substantial multicollinearity, with VIF values ranging from 1.29 to 2.72 and tolerance values ranging from 0.367 to 0.775. The Ramsey RESET test did not indicate any relevant model misspecification for either short-term surgical fear (p=0.676) or long-term surgical fear (p=0.514). The maximum Cook's distance was 0.092 for the short-term model and 0.049 for the long-term model, indicating that there were no highly influential observations. The Breusch-Pagan test did not indicate heteroscedasticity in the short-term model (p=0.273), but did indicate heteroscedasticity in the long-term model (p=0.036); therefore, HC3 robust standard errors were used for both models.

The multivariable model for short-term surgical fear explained 37.6% of the variance (R^2=0.376; adjusted R^2=0.333). Higher HADS-A scores were independently associated with higher short-term surgical fear (B=1.167, β=0.447, HC3 SE=0.259, p<0.001, 95% CI 0.655 to 1.678). Male sex was associated with lower short-term surgical fear scores compared with female sex (B=-4.879, β=-

Table 2. Comparison of continuous clinical and psychological variables between the trauma and elective surgery groups

Variable	**Trauma (n=64), median (Q1-Q3)**	**Elective (n=78), median (Q1-Q3)**	**p**	**Rank-biserial r**	**95% CI**
Age	56.5 (27.25-74)	61.5 (42-67.25)	0.925	-0.009	-0.225 to 0.199
Preoperative length of stay	3 (2-3)	1 (1-1)	<0.001	0.838	0.743 to 0.922
Depression score	9 (6-11)	8 (5-10)	0.083	0.168	-0.022 to 0.359
Anxiety score	9 (7-11.75)	8 (6-11.25)	0.163	0.136	-0.057 to 0.323
Total HADS score	19 (11-22.75)	17 (12-21)	0.161	0.137	-0.053 to 0.322
Short-term fear	23 (12-24)	17 (10-25)	0.175	0.132	-0.051 to 0.324
Long-term fear	20.5 (10.5-27)	16.5 (11.5-24)	0.112	0.155	-0.039 to 0.346
Total fear score	44 (28.25-50)	34 (18-47)	0.095	0.163	-0.030 to 0.354
VAS score	7 (6-9)	6 (4-7)	0.001	0.322	0.147 to 0.498

Group comparisons were performed using the Mann-Whitney U test. Rank-biserial correlation was used as the effect-size measure; positive values indicate higher ranks in the trauma group. CIs are bootstrap 95% confidence intervals. HADS: Hospital anxiety and depression scale, VAS: Visual analog scale, CI: Confidence interval

0.247, HC3 SE=1.699, p=0.005, 95% CI -8.239 to -1.518), indicating higher adjusted short-term fear among women. Previous surgery, trauma/elective status, age, marital status, preoperative length of stay, HADS-D, and VAS were not significantly associated with short-term fear (Table 4).

The multivariable model for long-term surgical fear explained 31.9% of the variance (R^2=0.319; adjusted R^2=0.272). Higher HADS-A scores (B=1.118, β=0.371, HC3 SE=0.337, p=0.001, 95% CI 0.451 to 1.785) and higher HADS-D scores (B=0.651, β=0.228, HC3 SE=0.289, p=0.026, 95% CI 0.080 to 1.223) were independently associated with increased long-term surgical fear. Trauma/elective status, sex, previous surgery, age, marital status, preoperative length of stay, and VAS were not significantly associated with long-term fear (Table 5).

Table 3. Correlations between age, psychological measures, surgical fear dimensions, and pain severity

Variable	1	2	3	4	5	6
1. Age	-					
2. Depression score	0.032	-				
3. Anxiety score	0.160	0.574**	-			
4. Short-term fear	0.206*	0.368**	0.480**	-		
5. Long-term fear	0.076	0.435**	0.476**	0.739**	-	
6. VAS score	0.226*	0.190*	0.322**	0.198*	0.205*	-

Data are Spearman's rho (ρ) coefficients. Significance was determined after Benjamini-Hochberg false discovery rate adjustment across the 15 unique pairwise correlations. *q<0.05; **q<0.01. VAS: Visual analog scale

Table 4. Multivariable linear regression analysis of factors associated with short-term surgical fear

Variable	B	HC3 robust SE	β	p	95% CI
Age	0.003	0.055	0.006	0.957	-0.107 to 0.112
Male sex vs. female	-4.879	1.699	-0.247	0.005	-8.239 to -1.518
Married vs. single	-0.851	2.248	-0.042	0.706	-5.298 to 3.596
Previous surgery ≥1 vs. none	2.732	1.731	0.134	0.117	-0.692 to 6.156
Trauma vs. elective	3.940	2.478	0.204	0.114	-0.963 to 8.842
Preoperative length of stay	-1.245	0.908	-0.141	0.172	-3.040 to 0.550
HADS-D	0.204	0.257	0.082	0.429	-0.305 to 0.712
HADS-A	1.167	0.259	0.447	<0.001	0.655 to 1.678
VAS	-0.114	0.380	-0.028	0.765	-0.865 to 0.638

R^2=0.376; adjusted R^2=0.333. B: Unstandardized regression coefficient, β: Standardized regression coefficient, HC3 SE: Heteroscedasticity-consistent robust standard error, CI: Confidence interval, HADS-D: Hospital anxiety and depression scale-depression, HADS-A: Hospital anxiety and depression scale-anxiety, VAS: Visual analog scale. Reference categories: female sex, single marital status, no previous surgery, and elective surgery

Table 5. Multivariable linear regression analysis of factors associated with long-term surgical fear

Variable	B	HC3 robust SE	β	p	95% CI
Age	-0.053	0.066	-0.098	0.422	-0.183 to 0.077
Male sex vs. female	-0.366	1.881	-0.016	0.846	-4.086 to 3.354
Married vs. single	2.008	2.861	0.086	0.484	-3.652 to 7.667
Previous surgery ≥1 vs. none	2.779	2.080	0.118	0.184	-1.336 to 6.895
Trauma vs. elective	2.598	2.974	0.117	0.384	-3.284 to 8.480
Preoperative length of stay	-0.332	1.141	-0.032	0.772	-2.589 to 1.925
HADS-D	0.651	0.289	0.228	0.026	0.080 to 1.223
HADS-A	1.118	0.337	0.371	0.001	0.451 to 1.785
VAS	-0.071	0.400	-0.015	0.859	-0.863 to 0.720

R^2=0.319; adjusted R^2=0.272. B: Unstandardized regression coefficient, β: Standardized regression coefficient, HC3 SE: Heteroscedasticity-consistent robust standard error, CI: Confidence interval, HADS-D: Hospital anxiety and depression scale-depression, HADS-A: Hospital anxiety and depression scale-anxiety, VAS: Visual analog scale. Reference categories: female sex, single marital status, no previous surgery, and elective surgery

Sensitivity analyses that additionally adjusted for ASA classification and caregiver status did not materially alter the principal findings. Trauma/elective status remained not significantly associated with either short-term fear (B=3.616, p=0.172) or long-term fear (B=1.365, p=0.672). Female sex remained associated with higher short-term fear (p=0.005), and HADS-A remained associated with short-term fear (p<0.001). For long-term fear, HADS-A (p=0.008) and HADS-D (p=0.014) remained statistically significant. Previous surgical history was not significant in either sensitivity model (p=0.115 and p=0.152, respectively).

DISCUSSION

The principal finding of this study was that there were no statistically significant differences between patients undergoing trauma-related and those undergoing elective orthopedic surgery in short-term, long-term, or total preoperative surgical fear. In contrast, patients in the trauma group had substantially longer preoperative hospital stays and higher preoperative pain scores. The multivariable analyses further showed that HADS-A was independently associated with both short- and long-term surgical fear; that female sex was independently associated with greater short-term fear; and that HADS-D was independently associated with greater long-term fear. Trauma/elective status was not independently associated with either fear dimension.

Importantly, the absence of statistically significant between-group differences should not be interpreted as evidence that trauma and elective patients are equivalent with respect to surgical fear. The study was designed for a conventional two-group comparison and did not include a prespecified equivalence margin. The small rank-biserial effect sizes for short-term, long-term, and total surgical fear, along with their CIs, provide information on the magnitude and precision of the observed differences, but do not establish formal equivalence.

The demographic and clinical profiles of the trauma and elective groups differed substantially. Orthopedic trauma populations encompass a broad spectrum of injury mechanisms, age groups, and levels of systemic burden,[12-15] whereas elective orthopedic populations commonly include older patients undergoing treatment for chronic degenerative conditions.[16,17] In this cohort, there was no statistically significant age difference between groups, which may reflect the heterogeneous composition of the trauma group and the inclusion of both younger and older patients.

The two groups also differed in marital status, number of children, caregiver status, previous hospitalization, previous surgery, and ASA classification. Caregiver availability may differ between unplanned trauma-related admissions and planned elective surgeries because patients undergoing elective procedures have more opportunity to organize social support before hospitalization. Caregiver burden and the contributions of spouses and family members have previously been described in patients undergoing treatment for hip fracture and in those undergoing elective joint replacement.[18-20] These baseline differences highlight the clinical heterogeneity of the two groups and support the use of adjusted analyses and sensitivity analyses, rather than interpretations based solely on crude group comparisons.

ASA distributions also differed markedly between groups. This difference may reflect variation in underlying systemic disease and case mix rather than the trauma/elective classification alone. ASA status has been associated with outcomes in orthopedic populations, including patients with hip fracture.[21] Nevertheless, additional adjustment for ASA classification and caregiver status did not materially change the associations observed in the primary multivariable analyses.

Surgical fear is a multidimensional psychological response that reflects concerns regarding the immediate surgical process, recovery, and longer-term outcomes. Studies using the SFQ in different surgical populations have demonstrated substantial variation in fear according to patient characteristics, psychological burden, and surgical context.[4,22-24] The present findings suggest that simply categorizing orthopedic surgery as trauma-related or elective may provide limited information regarding an individual patient's preoperative fear. Despite greater pain and longer hospitalization in the trauma group, trauma/elective status was not significantly associated with either fear dimension after adjustment.

HADS-A showed the strongest independent association with short-term surgical fear and was also independently associated with long-term surgical fear. This finding is consistent with previous evidence showing close relationships between preoperative anxiety, fear, worry, and perioperative psychological burden.[6,7,24] Although anxiety and surgical fear are related constructs, the SFQ specifically captures fear directed toward surgical and recovery-related consequences, whereas HADS-A measures broader anxiety symptoms. Their independent clinical assessment may, therefore, provide complementary information.

Depressive symptoms showed a different pattern. HADS-D was not independently associated with short-term fear but was associated with long-term surgical fear. Long-term SFQ items focus more strongly on recovery, functional outcomes, and possible adverse consequences after surgery. Cognitive models of depression emphasize negative expectations, pessimistic appraisal, and increased focus on unfavorable future outcomes,[25] which may partly explain why depressive symptoms were more strongly associated with long-term than immediate surgical concerns. However, because both variables were measured at the same preoperative time point, the present study cannot establish directionality or causality.

Preoperative length of hospital stay differed markedly between the groups but was not independently associated with either dimension of surgical fear. Previous studies have shown that waiting periods immediately before surgery or in the operating

room may influence anxiety and surgical fear.[26] The variable examined in the present study represented the hospital stay from admission to surgery, rather than the waiting time in the operating area. These measures therefore reflect different aspects of the perioperative process. In elective arthroplasty pathways, changes in perioperative organization may also considerably shorten hospitalization.[27] A more detailed evaluation of waiting-related variables may be useful in prospective studies.

Surgical history was also not independently associated with short- or long-term fear when previous operations were analyzed as a clinically interpretable binary variable. Prior surgical exposure may have heterogeneous psychological effects. Familiarity with the surgical process may reduce uncertainty in some patients, whereas negative previous experiences may increase fear in others.[10,22] The present data do not support treating previous surgery itself as an independent marker of increased surgical fear.

Female sex was independently associated with greater short-term, but not long-term, surgical fear. Sex-related differences in preoperative fear have been reported in surgical populations, although the magnitude and dimensions of these associations vary across studies.[10,22] In the present study, the association was specific to short-term concerns, suggesting that sex-related differences may be more apparent for immediate fears surrounding surgery and anesthesia than for longer-term recovery-related concerns.

Pain showed weak bivariate associations with both fear dimensions and anxiety and depressive symptoms. However, VAS score was not independently associated with either short- or long-term fear in the multivariable models. Pain is an important component of the orthopedic perioperative experience and has been associated with psychological and postoperative outcomes,[28] but the present findings indicate that preoperative anxiety and depressive symptoms showed stronger adjusted associations with surgical fear than pain intensity in the multivariable models.

These findings have practical implications for preoperative assessment. Rather than relying solely on trauma or elective indications, pain severity, or previous surgical exposure to identify patients with greater surgical fear, clinicians may benefit from directly considering psychological symptoms. Structured preoperative education has been associated with reductions in anxiety in patients undergoing elective hip and knee arthroplasty,[29] and individualized education and psychological support may help reduce uncertainty and improve preparedness for surgery.

A multidisciplinary approach may be particularly relevant in orthopedic care, where physical injury, pain, function, social support, and mental health frequently interact. Reviews of orthopedic trauma care have emphasized the importance of integrating psychosocial assessment and support with conventional orthopedic treatment.[3,30] The present findings suggest that such assessment should not be limited to trauma patients but should be considered based on individual psychological needs across both trauma-related and elective surgical populations.

Study Limitations

This study has several limitations. Its single-center, cross-sectional design limits generalizability and precludes establishing causal or temporal relationships. The trauma and elective groups were clinically heterogeneous, and residual confounding may remain despite multivariable and sensitivity analyses. Detailed trauma characteristics, including mechanism and injury severity, were not available. Emergency cases were excluded, which may limit the applicability of the findings to more severe trauma populations. In addition, SFQ, HADS, and VAS were self-reported and assessed at a single preoperative time point. The study was powered to detect a medium between-group effect; therefore, smaller clinically relevant differences cannot be excluded. Other potentially important factors, including previous surgical experiences, social support, perioperative information needs, personality characteristics, and patient–physician communication, were not comprehensively assessed.

CONCLUSION

In this cross-sectional study, no statistically significant differences were detected in short-term, long-term, or total preoperative surgical fear between patients undergoing trauma-related orthopedic surgery and those undergoing elective orthopedic surgery. These findings should not be interpreted as demonstrating equivalence between the groups. Higher levels of anxiety symptoms were independently associated with both short- and long-term surgical fear. Female sex and depressive symptoms were independently associated with greater short-term and long-term fear, respectively. Trauma versus elective surgical status, surgical history, preoperative pain, and preoperative length of hospital stay were not independently associated with either of the fear dimensions after multivariable adjustment. These findings indicate that individual psychological characteristics may provide clinically relevant information beyond the surgical indication alone. Preoperative assessment of anxiety, depressive symptoms, and surgical fear may help identify patients who could benefit from individualized psychological support or preoperative education. Prospective longitudinal studies are required to determine whether these associations influence postoperative recovery and whether targeted interventions can improve clinical outcomes.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the University of Health Sciences Türkiye, Koşuyolu Yüksek İhtisas Training and Research Hospital Ethics Committee (decision no: 2026/04/1386, date: 24.02.2026).

Informed Consent: The study was conducted in accordance with the Declaration of Helsinki and written informed consent was

obtained from all participants prior to participation in the study. All study data were anonymized and used solely for scientific purposes.

Footnotes

Authorship Contributions

Surgical and Medical Practices: H.Ç.K., Concept: H.Ç.K., B.B.K., Design: H.Ç.K., Data Collection or Processing: H.Ç.K., B.B.K., A.S.K., Analysis or Interpretation: B.B.K., O.T., B.Y., Literature Search: H.Ç.K., B.B.K., Writing: H.Ç.K., B.B.K.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Kumar S, Verma V, Kushwaha U, Calvello Hynes EJ, Arya A, Agarwal A. Prevalence and association of depression in in-patient orthopaedic trauma patients: a single centre study in India. J Clin Orthop Trauma. 2020;11:S573-7.
2. Yang Y, Tang TT, Chen MR, Xiang MY, Li LL, Hou XL. Prevalence and association of anxiety and depression among orthopaedic trauma inpatients: a retrospective analysis of 1994 cases. J Orthop Surg Res. 2020;15:587.
3. Scott S, Brameier DT, Tryggedsson I, et al. Prevalence, resources, provider insights, and outcomes: a review of patient mental health in orthopaedic trauma. J Orthop Surg Res. 2024;19:538.
4. Theunissen M, Peters ML, Schouten EG, et al. Validation of the surgical fear questionnaire in adult patients waiting for elective surgery. PLoS One. 2014;9:e100225.
5. Bağdigen M, Karaman Özlü Z. Validation of the Turkish version of the surgical fear questionnaire. J Perianesth Nurs. 2018;33:708-14.
6. Shebl MA, Toraih E, Shebl M, et al. Preoperative anxiety and its impact on surgical outcomes: a systematic review and meta-analysis. J Clin Transl Sci. 2025;9:e33.
7. Ayyadhah Alanazi A. Reducing anxiety in preoperative patients: a systematic review. Br J Nurs. 2014;23:387-93.
8. Saray Kılıç H, İbrahimoğlu Ö, Andıç K. Impact of preoperative surgical fear and anxiety on sleep quality and recovery outcomes after arthroplasty. J Educ Res Nurs. 2025;22:233-8.
9. Kılınç T, Karaman Özlü Z. The determination of the relationship between fear of surgery, sleep, and insomnia in patients scheduled for elective surgery. SBÜHD. 2023;5:205-12.
10. Bez H, Erturhan Türk K. Determination of preoperative surgical fear in orthopedic surgery patients: a descriptive cross-sectional study. Cumhuriyet Nursing Journal. 2024;8:118-24.
11. Aydemir Ö, Güvenir T, Küey L, Kültür S. Hastane anksiyete ve depresyon ölçeği Türkçe formunun geçerlilik ve güvenilirliği. Türk Psikiyatri Dergisi. 1997;8:280-7.
12. Ranson R, Esper GW, Covell N, et al. Age is not just a number: the intersection of age, orthopedic injuries, and worsening outcomes following low-energy falls. J Geriatr Phys Ther. 2025;48:24-31.
13. Taylor A, Young A. Epidemiology of orthopaedic trauma admissions over one year in a district general hospital in England. Open Orthop J. 2015;9:191-3.
14. Laney W, Naicker D, Milner B, Omar S. Analysis of orthopaedic injuries in CT pan scans of polytrauma patients at a quaternary academic hospital. SA Orthopaedic Journal. 2023;22:76-81.
15. Okkaoglu MC, Özdemir M, Şeşen H, Taşkesen A, Demirkale İ, Altay M. Retrospective analysis of epidemiological characteristics of elderly hip fractures. AATD. 2016;1:17-23.
16. Yılmaz Ferhatoğlu S, Dönmez NF. The effect of nutritional status on length of hospital stay in adult patients undergoing elective orthopedic surgery: a prospective analysis. Med Bull Haseki. 2020;58:228-33.
17. Kappenschneider T, Maderbacher G, Weber M, et al. Special orthopaedic geriatrics (SOG) - a new multiprofessional care model for elderly patients in elective orthopaedic surgery: a study protocol for a prospective randomized controlled trial of a multimodal intervention in frail patients with hip and knee replacement. BMC Musculoskelet Disord. 2022;23:1079.
18. Xiao P, Zhou Y. Factors associated with the burden of family caregivers of elderly patients with femoral neck fracture: a cross-sectional study. J Orthop Surg Res. 2020;15:234.
19. Stark ÅJ, Salanterä S, Sigurdardottir AK, Valkeapää K, Bachrach Lindström M. Spouse-related factors associated with quality of recovery of patients after hip or knee replacement – a Nordic perspective. Int J Orthop Trauma Nurs. 2016;23:32-46.
20. Çeçen Çamlı D, Kara H. Evaluation of care burden and quality of life of the caregivers in elderly patients in surgical clinics. Sağ Aka Derg. 2025;12:25-32.
21. Badghish EM, Alsini AY, Alnemari IH, Alfaify AM. Factors associated with one-year mortality after hip fractures: a five-year retrospective cohort study. J Musculoskelet Surg Res. 2025;9:329-34.
22. Çağlar S, Karaman Özlü Z. Detection of surgical fear degree of patients on X, Y, Z generation. Karya J Health Sci. 2022;3:69-77.
23. Karačić A, Theunissen M, Sović S, Sever M, Bakula B, Semanjski K. Validation of the croatian version of the surgical fear questionnaire (SFQ) in adult patients waiting for elective surgery. Acta Clin Croat. 2023;62:153-61.
24. Nielsen AH, Gamst-Jensen H, Buch B, et al. Patient-reported surgical fear and degree of worry on the day of surgery: correlation, predictors, and implications for patient-centered care. A mixed-methods study. J Perianesth Nurs. 2026;41:754-61.
25. LeMoult J, Gotlib IH. Depression: a cognitive perspective. Clin Psychol Rev. 2019;69:51-66.
26. Koyuncu A, Yakıt E. The effect of waiting time on the surgical stretcher in the operating room for urological surgery on anxiety and surgical fear: a comparative cross-sectional study. J Perianesth Nurs. 2025;40:1535-40.
27. den Hartog YM, Mathijssen NM, Vehmeijer SB. Reduced length of hospital stay after the introduction of a rapid recovery protocol for primary THA procedures. Acta Orthop. 2013;84:444-7.
28. Riecke J, Zerth SF, Schubert AK, et al. Risk factors and protective factors of acute postoperative pain: an observational study at a German university hospital with cross-sectional and longitudinal inpatient data. BMJ Open. 2023;13:e069977.
29. Lloyd A, Verma S, Crane L. Preoperative education for reducing patient anxiety in elective hip and knee arthroplasty – a systematic review and meta-analysis. Perioper Med (Lond). 2026;15:11.
30. Andsoy II. Care and current approaches in orthopedic trauma patients. TOTBİD Dergisi. 2025;24:148-57.

DOI: 10.4274/forbes.galenos.2026.57983
Forbes J Med 2026;7:165-70

Original Article

Functional Outcomes of Reduction Mammoplasty Assessed by the Neck Disability Index: A Retrospective Study with Long-Term Follow-Up

Uğur HOROZ[1], Uğur GÜNGÖR DEMİR[2], Ali Nail DEMİR[3]

[1]Private Clinic, Clinic of Plastic, Reconstructive and Aesthetic Surgery, İstanbul, Türkiye
[2]University of Health Sciences Türkiye, Mersin City Hospital, Clinic of Physical Medicine and Rehabilitation, Mersin, Türkiye
[3]University of Health Sciences Türkiye, Mersin City Hospital, Clinic of Rheumatology, Mersin, Türkiye

Cite as: Horoz U, Demir UG, Demir AN. Functional outcomes of reduction mammoplasty assessed by the neck disability index: a retrospective study with long-term follow-up. Forbes J Med. 2026;7:165-70

ABSTRACT

Objective: To evaluate the effect of reduction mammoplasty on neck-related disability using the neck disability index (NDI) and to assess its utility as a functional outcome measure in women with symptomatic macromastia.

Methods: This retrospective cohort study included 107 consecutive women who underwent reduction mammoplasty for symptomatic macromastia between January 2019 and December 2023. Demographic characteristics, operative details, postoperative complications, and NDI scores were analyzed. The mean age was 44 years, the mean body mass index was 27.3 kg/m^2, the mean resection weight was 1,283 g (range, 450-4,100 g), and the mean follow-up was 28 months (range, 4-48 months). Preoperative and postoperative NDI scores were compared.

Results: The mean NDI score decreased significantly from 39.7 preoperatively to 8.6 postoperatively, representing a 78.3% reduction in neck-related disability ($p<0.001$). Minor wound dehiscence occurred in 11 patients, while 6 patients required secondary wound closure. 1 patient developed a hematoma, 1 experienced partial nipple-areola complex necrosis, and 4 underwent revision reduction mammoplasty.

Conclusion: Reduction mammoplasty provides substantial and sustained improvement in neck-related disability among women with symptomatic macromastia. The marked reduction in NDI scores supports its role as a functional reconstructive procedure. The NDI is a practical patient-reported outcome measure for assessing functional recovery after breast reduction surgery.

Keywords: Reduction mammoplasty, macromastia, neck disability index, neck pain, breast hypertrophy, functional outcome

Received: 04.06.2026
Accepted: 09.10.2026
Publication Date: 12.10.2026

Corresponding Author:
Uğur HOROZ
Private Clinic, Clinic of Plastic, Reconstructive and Aesthetic Surgery, İstanbul, Türkiye
ugur_horoz@hotmail.com
ORCID ID: 0000-0003-4604-8934

INTRODUCTION

Macromastia is a common condition associated with substantial physical and psychosocial morbidity. Women with symptomatic breast hypertrophy frequently experience chronic neck, shoulder, and back pain, shoulder grooving, intertriginous skin disorders, physical limitations, and impaired quality of life. These symptoms often interfere with daily activities, occupational performance, physical exercise, and social participation.[1,2]

The pathophysiology of musculoskeletal symptoms associated with macromastia is multifactorial. Excessive breast weight shifts the body's center of gravity anteriorly, leading to compensatory postural adaptations and increased mechanical loading of the cervical and thoracic spine.[3] Previous studies have demonstrated associations between breast hypertrophy and thoracic kyphosis, altered spinal alignment, postural imbalance, and increased biomechanical stress.[4-11] Over time, these biomechanical alterations may contribute to chronic neck and back pain, progressive postural deterioration, and degenerative changes of the spine.[6-10]

Reduction mammoplasty has gained wide acceptance as the treatment of choice for symptomatic macromastia and has proven to be effective in relieving musculoskeletal symptoms and improving physical function and health-related quality of life.[2,3,6,8,10,11] Objective biomechanical investigations have further demonstrated reductions in spinal loading and improvements in functional performance following breast reduction surgery, reinforcing the reconstructive rather than purely aesthetic role of the procedure.[3]

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

Although numerous studies have evaluated pain relief and patient satisfaction after reduction mammoplasty, relatively few have assessed neck-specific functional impairment using validated patient-reported outcome measures.[12-14] The neck disability index (NDI), originally developed by Vernon and Mior,[12] is one of the most widely accepted and extensively validated instruments for evaluating functional disability related to cervical musculoskeletal disorders.[13,14] Unlike pain intensity scales, the NDI provides a comprehensive assessment of how neck symptoms affect activities of daily living and functional performance. Despite its widespread use in spine research, its application in reduction mammoplasty remains limited.

Therefore, the aim of the present study was to assess changes in neck disability in patients who underwent reduction mammoplasty using the NDI, and to explore long-term functional outcomes through interdisciplinary collaboration between the departments of plastic surgery and physical medicine and rehabilitation.

METHODS

Study Design and Patients

The retrospective cohort study involves consecutive patients who underwent reduction mammoplasty for symptomatic macromastia between 2019 and 2023. Medical records of patients who completed preoperative and postoperative diagnostic tests were reviewed. All participants who met inclusion criteria and did not meet any of the exclusion criteria were enrolled in the study. There were 107 female patients who qualified for the research. Those with incomplete sets of clinical records or absent evaluations of the NDI were excluded from the analysis.

The study followed the ethical guidelines of the Declaration of Helsinki. The anonymity of the participants was ensured, and all patient data were de-identified before analysis. The confidentiality of patient information was maintained during the study.

Multidisciplinary Evaluation

All patients were evaluated collaboratively by the departments of plastic surgery and physical medicine and rehabilitation. The patients were assessed clinically for neck pain, shoulder pain, upper back pain, postural complaints, and limitations.

Data Collection

Demographic and clinical variables were obtained from institutional records: age, body mass index (BMI), resection specimen weight, length of postoperative follow-up, and postoperative complications. The mean age of the patients is 44 years, and the mean BMI is 27.3 kg/m^2. The operative data included the total bilateral amount of resected breast tissue. The mean combined bilateral resection weight was 1,283 g, ranging from 450 g to 4,100 g.

Outcome Measure

Neck disability was assessed using the NDI, a validated patient-reported outcome measure widely used to assess patient-reported outcomes related to cervical musculoskeletal disorders.

The validated Turkish version of the NDI was used as a patient-reported outcome measure. The NDI comprises 10 items, each scored on a scale from 0 to 5, with a total score ranging from 0 to 50 that represents the degree of limitation in activities of daily living caused by neck complaints. The preoperative NDI scores were obtained as part of the standard preoperative assessment. The NDI scores used for analysis were obtained at the patient's most recent clinic visit, which was defined as the study endpoint

Follow-up

The mean postoperative follow-up period was 28 months (range, 4-48 months).

Statistical Analysis

Statistical analysis was carried out using IBM SPSS Statistics for Windows, Version 29.0 (IBM Corp., Armonk, NY, USA). For continuous variables, means and ranges are presented. Paired statistical analysis was used to compare the NDI values before and after the intervention. A p value of <0.05 was considered statistically significant.

Ethical Approval and Informed Consent

The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. The research was carried out in a private clinic without an institutional ethics committee. Therefore, ethical approval was obtained from a Clinical Research Ethics Committee located in Ankara Bilkent City Hospital (approval number: E2-26-15707; date: 22 July 2026). Written informed consent for the surgical procedures was obtained from all patients. Additionally, written consent for the scientific use and publication of clinical photographs was obtained from all patients whose images appear in this manuscript. Patient records were retrospectively reviewed, and all data were anonymized before analysis.

RESULTS

Patient Characteristics

A total of 107 patients underwent reduction mammoplasty for symptomatic macromastia during the study period (Figures 1 and 2).

Their mean age was 44 years, and their mean BMI was 27.3 kg/m^2. The mean combined bilateral resection weight was 1,283 g (range, 450-4,100 g). Their mean follow-up was 28 months (range, 4-48 months). Patient demographics and clinical characteristics are detailed in Table 1.

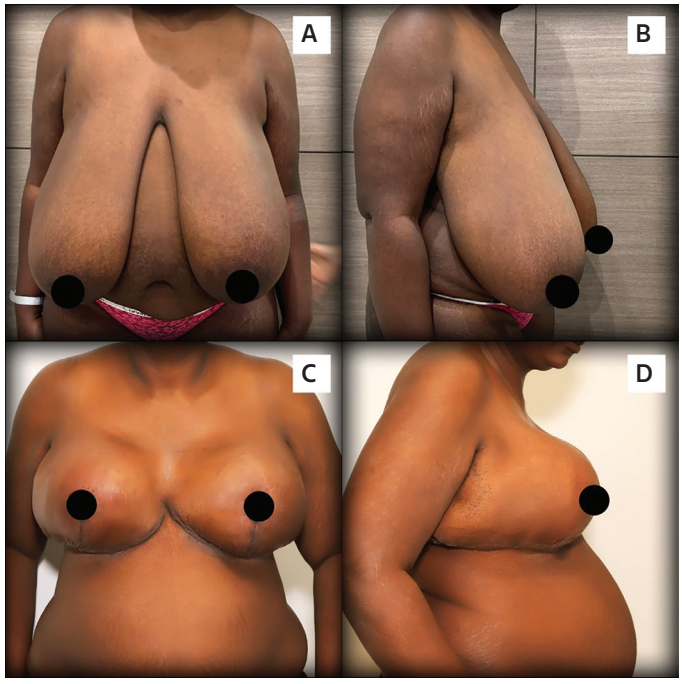


Figure 1. Representative preoperative and postoperative clinical photographs of a 46-year-old woman. (A, B) Preoperative anterior and right lateral views demonstrating symptomatic macromastia and severe breast ptosis. (C, D) Anterior and right lateral views obtained 9 months after reduction mammoplasty, demonstrating marked improvement in breast size, contour, symmetry, and nipple-areola complex position

Table 1. Demographic characteristics, follow-up duration, body mass index (BMI), and combined bilateral breast tissue resection weight of the 107 women included in the study

Variable	Value
Number of patients	107
Mean age (years)	44
Mean BMI (kg/m^2)	27.3
Mean follow-up (months)	28
Follow-up range (months)	4-48
Mean combined bilateral resection weight (g)	1,283
Combined bilateral resection range (g)	450-4,100

Table 2. Preoperative and postoperative neck disability index (NDI) scores demonstrating the effect of reduction mammoplasty on neck-related functional disability

Variable	Value
Preoperative NDI	39.7
Postoperative NDI	8.6
Absolute improvement	31.1
Relative improvement (%)	78.3
p value	<0.001

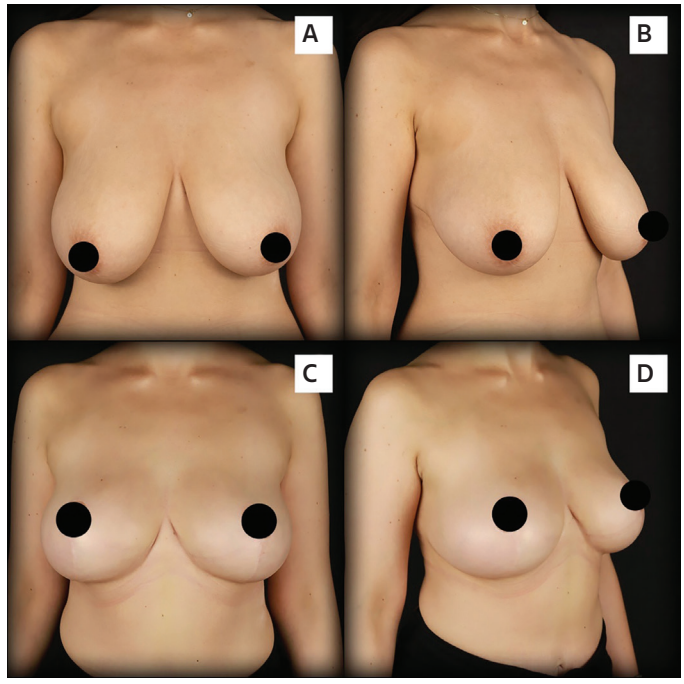


Figure 2. Representative preoperative and postoperative clinical photographs of a 28-year-old woman. (A, B) Preoperative anterior and right oblique views demonstrating symptomatic macromastia and breast ptosis. (C, D) Anterior and right oblique views obtained 2 years after reduction mammoplasty, demonstrating sustained improvement in breast size, contour, projection, symmetry, and nipple-areola complex position

Neck Disability Index Outcomes

The mean preoperative and postoperative NDI scores were 39.7 and 8.6, respectively.

The mean score improved by 31.1 points, representing a 78.3% improvement in neck disability.

Comparison of preoperative and postoperative scores demonstrated a statistically significant difference ($p<0.001$). The detailed results of the NDI are presented in Table 2.

This result indicates a substantial reduction in neck-related disability and improvement in functional status following reduction mammoplasty. The significant decrease in the NDI score after the procedure is shown in Figure 3.

Postoperative Complications

Postoperative wound dehiscence, managed conservatively with local wound care, occurred in 11 patients (10.3%).

Six patients (5.6%) developed wound dehiscence requiring secondary suturing under local anesthesia. One patient (0.9%) developed a postoperative hematoma requiring intervention. Partial nipple-areola complex necrosis occurred in one patient (0.9%).

No cases of total nipple-areola complex loss were observed. Most complications were minor wound-healing problems managed conservatively or with secondary suturing. Postoperative complications are summarized in Table 3.

Secondary/Revision Procedures

Four patients (3.7%) underwent secondary revision reduction mammoplasty because they considered the postoperative breast volume to be larger than desired. These procedures were classified as secondary/revision procedures rather than postoperative complications.

No cases of total nipple-areola complex loss were observed.

DISCUSSION

The present study demonstrated a substantial and sustained improvement in neck-related disability following reduction mammoplasty. The mean NDI score decreased from 39.7 preoperatively to 8.6 postoperatively, representing a 78.3% reduction in disability over a mean follow-up of 28 months. These findings indicate that reduction mammoplasty provides not only effective symptom relief but also meaningful long-term functional recovery in patients with symptomatic macromastia.

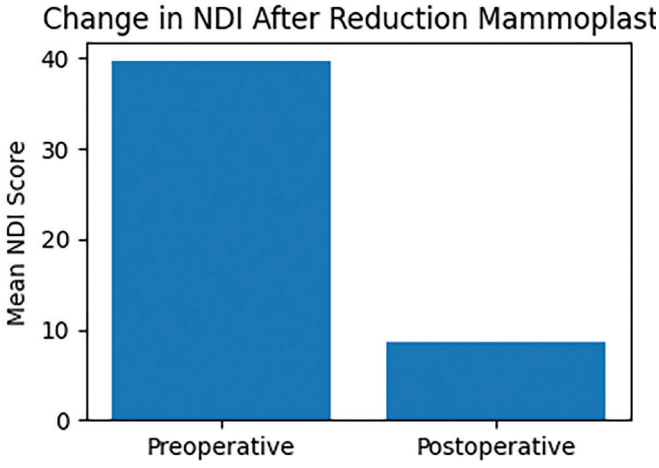


Figure 3. Change in neck disability index (NDI) after reduction mammoplasty. Comparison of mean preoperative and postoperative NDI scores following reduction mammoplasty. The mean NDI score decreased from 39.7 before surgery to 8.6 after surgery, corresponding to a 78.3% reduction in neck-related disability ($p<0.001$), indicating significant functional improvement

Table 3. Postoperative complications following reduction mammoplasty

Complication	n (%)
Conservative wound dehiscence	11 (10.3)
Dehiscence requiring suturing	6 (5.6)
Hematoma	1 (0.9)
Partial NAC necrosis	1 (0.9)

NAC: Nipple-areola complex

Musculoskeletal complaints represent the primary indication for reduction mammoplasty in women with symptomatic breast hypertrophy. Netscher et al.[7] reported that 81% of patients experienced neck and back pain, 77% experienced shoulder pain, and nearly half experienced limitations in daily activities attributable to breast size. Similarly, Brown et al.[8] demonstrated significant postoperative improvements in neck, shoulder, and back pain, and all patients reported increased participation in physical activity. Our findings are consistent with these reports and further demonstrate that pain relief is accompanied by a marked improvement in neck-specific functional disability.

Several biomechanical mechanisms may explain these clinical improvements. Excessive breast weight shifts the body's center of gravity anteriorly, increasing mechanical loading of the cervical and thoracic spine and resulting in compensatory postural adaptations. Letterman and Schurter were among the first to suggest that excessive breast weight adversely affects spinal mechanics and cervical posture.[11] More recently, Lapid et al.[9] and Michalik et al.[10] demonstrated measurable changes in thoracic kyphosis, spinal alignment, and postural balance associated with increasing breast size. Although BMI also influences spinal posture, these studies support the conclusion that breast hypertrophy independently contributes substantially to musculoskeletal dysfunction.[9,10]

Objective biomechanical evidence further supports the functional benefits of reduction mammoplasty. Foreman et al.[3] demonstrated a 35% reduction in low-back compressive forces together with a 76% improvement in functional rating index scores following surgery. While their study primarily evaluated lumbar biomechanics, our findings extend this evidence by demonstrating a substantial improvement in cervical functional disability, as measured by the NDI. This distinction is clinically important because the cervical spine is frequently the primary source of symptoms in women with macromastia.

Recent clinical studies continue to reinforce the therapeutic value of reduction mammoplasty. Chinondira et al.[2] reported significant reductions in neck, shoulder, and back pain following surgery, emphasizing its effectiveness in improving functional status. Likewise, recent reviews have highlighted the cumulative effects of chronic spinal loading, postural compensation, and altered biomechanics as key contributors to persistent musculoskeletal symptoms in women with macromastia.[1]

One of the principal strengths of the present study is the use of the NDI as the primary outcome measure. The present study contributes to the limited literature that specifically evaluates neck-related functional disability after reduction mammoplasty, using a validated neck-specific patient-reported outcome measure.

Although reduction mammoplasty has consistently been shown to reduce pain and improve quality of life, relatively few studies have evaluated postoperative outcomes using validated neck-specific functional instruments. Unlike studies primarily focused on pain relief, patient satisfaction, or general quality-of-life outcomes, our study provides a standardized assessment of cervical functional recovery. The NDI is one of the most extensively validated patient-reported outcome measures for assessing cervical disability and offers a more comprehensive evaluation of postoperative functional recovery than pain intensity alone.[12-14] Furthermore, the multidisciplinary collaboration between the departments of plastic surgery and physical medicine and rehabilitation enabled a comprehensive functional assessment, thereby strengthening the clinical relevance of our findings.

The mean combined bilateral resection weight in our series was 1,283 g, reflecting a cohort with clinically significant macromastia. Previous investigations have suggested that symptom severity and postoperative improvement correlate more closely with functional impairment than with resection weight alone.[2,3,6] These findings support the concept that surgical decision-making should primarily be guided by symptom burden and functional disability rather than by the amount of breast tissue expected to be removed.

The safety profile observed in our cohort was consistent with those reported in previously published reduction mammoplasty series. Most complications were minor wound-healing problems successfully managed with conservative treatment or minor revision procedures; major complications were uncommon. These findings further support reduction mammoplasty as a safe and effective reconstructive procedure when performed in appropriately selected patients.[2,4-6]

Study Limitations

Several limitations should be acknowledged. First, the retrospective design may have introduced selection bias, particularly because the analysis was limited to patients who had complete preoperative and postoperative NDI assessments. Second, the absence of a control group limits causal inference. Third, postoperative NDI assessment was based on the most recent follow-up visit; the variability in follow-up duration (4-48 months) may have influenced the timing of the functional outcome assessment. Fourth, although the NDI provides a robust assessment of neck-related disability, additional validated breast-specific patient-reported outcome measures (such as the BREAST-Q), as well as generic quality-of-life questionnaires, were not routinely available for analysis. Finally, potential confounding effects of age, BMI, resection weight, and coexisting musculoskeletal conditions could not be fully evaluated within the retrospective study design. Nevertheless, the long-term follow-up and substantial magnitude of functional improvement support the clinical relevance of our findings.

Overall, the present findings support the concept that reduction mammoplasty should be regarded as a functional reconstructive procedure rather than merely an aesthetic intervention. In addition to improving breast appearance, the procedure provides substantial and sustained reductions in neck-related disability and improvements in functional capacity in women with symptomatic macromastia. These findings further emphasize the importance of considering symptom burden and functional impairment when evaluating candidates for reduction mammoplasty. Prospective studies incorporating standardized assessment intervals, breast-specific patient-reported outcome measures, and analyses of potential predictors of functional recovery are warranted to further clarify the long-term functional benefits of reduction mammoplasty.[2,12-14]

CONCLUSION

Reduction mammoplasty resulted in a significant and sustained improvement in neck-related disability among women with symptomatic macromastia, with a 78.3% reduction in NDI scores over a mean follow-up period of 28 months. These findings reinforce the role of reduction mammoplasty as a functional reconstructive procedure that provides substantial long-term benefits beyond aesthetic improvement.

The NDI proved to be a practical and valuable patient-reported outcome measure for assessing functional recovery after reduction mammoplasty. Its incorporation into routine clinical practice may facilitate standardized outcome assessment and provide objective evidence supporting the medical necessity of breast reduction surgery.

Furthermore, a multidisciplinary approach involving specialists in Plastic Surgery and Physical Medicine and Rehabilitation may enhance preoperative evaluation, postoperative follow-up, and assessment of functional outcomes. Prospective studies incorporating disease-specific and quality-of-life instruments are warranted to further validate the role of the NDI in evaluating long-term outcomes after reduction mammoplasty.

Ethics

Ethics Committee Approval: The study was conducted in accordance with the ethical principles of the Declaration of Helsinki. The research was carried out in a private clinic without an institutional ethics committee. Therefore, ethical approval was obtained from a Clinical Research Ethics Committee located in Ankara Bilkent City Hospital (approval number: E2-26-15707; date: 22 July 2026).

Informed Consent: Written informed consent for the surgical procedures was obtained from all patients. Additionally, written consent for the scientific use and publication of clinical photographs was obtained from all patients whose images appear in this manuscript. Patient records were retrospectively reviewed, and all data were anonymized before analysis.

Footnotes

Authorship Contributions

Surgical and Medical Practices: U.H., U.G.D., Concept: U.H., A.N.D., Design: U.G.D., A.N.D., Data Collection or Processing: U.H., A.N.D., Analysis or Interpretation: U.H., U.G.D., Literature Search: U.H., U.G.D., A.N.D., Writing: U.H., U.G.D.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. González-Mendoza DE, Rábago-Sánchez PP. Gigantomastia and physical affections: a mini-review. Rev Fac Med Univ Nac Auton Mex. 2026;14:138-44.
2. Chinondira BB, Adan AA, Nango'le FW. Reduction mammoplasty in the treatment of neck, back and shoulder pain. Pan Afr J Plast Reconstr Aesthet Surg. 2025;2:67-72.
3. Foreman KB, Dibble LE, Droge J, Carson R, Rockwell WB. The impact of breast reduction surgery on low-back compressive forces and function in individuals with macromastia. Plast Reconstr Surg. 2009;124:1393-9.
4. Spencer L, Briffa K. Breast size, thoracic kyphosis & thoracic spine pain - association & relevance of bra fitting in post-menopausal women: a correlational study. Chiropr Man Therap. 2013;21:20.
5. Chadbourne EB, Zhang S, Gordon MJ, et al. Clinical outcomes in reduction mammaplasty: a systematic review and meta-analysis of published studies. Mayo Clin Proc. 2001;76:503-10.
6. Blomqvist L, Brandberg Y. Three-year follow-up on clinical symptoms and health-related quality of life after reduction mammaplasty. Plast Reconstr Surg. 2004;114:49-54.
7. Netscher DT, Meade RA, Goodman CM, Brehm BJ, Friedman JD, Thornby J. Physical and psychosocial symptoms among 88 volunteer subjects compared with patients seeking plastic surgery procedures to the breast. Plast Reconstr Surg. 2000;105:2366-73.
8. Brown JR, Holton LH 3rd, Chung TL, Slezak S. Breast-feeding, self-exam, and exercise practices before and after reduction mammoplasty. Ann Plast Surg. 2008;61:375-9.
9. Lapid O, de Groof EJ, Corion LU, Smeulders MJ, van der Horst CM. The effect of breast hypertrophy on patient posture. Arch Plast Surg. 2013;40:559-63.
10. Michalik R, Kühlmann B, Wild M, et al. The effect of breast size on spinal posture. Aesthetic Plast Surg. 2024;48:1331-8.
11. Piper ML, Esserman LJ, Sbitany H, Peled AW. Outcomes following oncoplastic reduction mammoplasty: a systematic review. Ann Plast Surg. 2016;76 Suppl 3:S222-6.
12. Vernon H, Mior S. The neck disability index: a study of reliability and validity. J Manipulative Physiol Ther. 1991;14:409-15. Erratum in: J Manipulative Physiol Ther 1992;15:followi.
13. Vernon H. The neck disability index: state-of-the-art, 1991-2008. J Manipulative Physiol Ther. 2008;31:491-502.
14. Kazarian GS, Steinhaus ME, Iyer S, et al. Baseline neck disability index and patient-reported outcomes measurement information system physical function predict postoperative return to normal in cervical spine surgery. Int J Spine Surg. 2024;18:653-9.

DOI: 10.4274/forbes.galenos.2026.84666
Forbes J Med 2026;7:171-7

Original Article

Electrocardiographic Predictors of Early Atrial Fibrillation After Transcatheter Atrial Septal Defect Closure: The Roles of P-Wave Dispersion and P-Terminal Force in V1

Yusuf DEMİR[1], Ahmet Anıl BAŞKURT[1], Samet SEVİNÇ[2], Oktay ŞENÖZ[1]

[1]Bakırçay University Çiğli Training and Research Hospital, Department of Cardiology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, Mehmet Akif Ersoy Thoracic and Cardiovascular Surgery Training and Research Hospital, Clinic of Cardiology, İstanbul, Türkiye

Cite as: Demir Y, Başkurt AA, Sevinç S, Şenöz O. Electrocardiographic predictors of early atrial fibrillation after transcatheter atrial septal defect closure: the roles of P-wave dispersion and P-terminal force in V1. Forbes J Med. 2026;7:171-7

ABSTRACT

Objective: Early atrial fibrillation (AF) can complicate transcatheter atrial septal defect (ASD) closure, yet simple electrocardiographic predictors remain underused. We investigated whether baseline P-wave dispersion (PWDis) and P-terminal force in V1 (PTFV1) predicted AF within 48 hours post-procedure.

Methods: We conducted a retrospective, two-center cohort study including 555 consecutive adults (18-85 years) who underwent successful percutaneous secundum ASD closure between March 2015 and September 2024. Patients with prior AF, coronary artery disease, failed procedures, emergent surgery, in-hospital death within 48 hours, suboptimal records, or unavailable preprocedural ECGs were excluded. Baseline 12-lead ECGs (25 mm/s, 10 mm/mV; 0.05-100 Hz) were digitized and measured at 4× magnification. PWDis was calculated as the difference between maximum and minimum P-wave duration across 12 leads; PTFV1 was defined as the product of the duration and amplitude of the terminal negative P-wave component in V1 (μV·ms).

Results: Of 555 patients, 27 (4.9%) developed AF within 48 hours. Patients with AF were older (59.6±13.1 vs. 37.9±14.7 years, p<0.001) and more often NYHA class II (p<0.001). On ECG, PWDis and PTFV1 were significantly higher in the AF group (PWDis 42.8±13.9 vs. 36.4±11.8 ms, p=0.007; PTFV1 4325.9±730.4 vs. 3876.5±656.1 μV·ms, p=0.001). Both PWDis [odds ratio (OR) 1.036; 95% confidence interval (CI) 1.004-1.068; p=0.025] and PTFV1 (OR 1.086; 95% CI 1.027-1.149; p=0.004) independently predicted AF. ROC analysis identified PTFV1 >4,050 μV·ms and PWDis >36.5 ms as predictors of early AF.

Conclusion: Preprocedural PWDis and PTFV1 independently predict early AF after ASD device closure. These ubiquitous, low-cost ECG indices may help risk-stratify candidates and tailor peri-procedural monitoring and prophylaxis strategies.

Keywords: Atrial septal defect, atrial fibrillation, P-wave dispersion, P-terminal force, noninvasive electrocardiology, transcatheter closure

Received: 18.04.2026
Accepted: 22.09.2026
Publication Date: 14.10.2026

Corresponding Author:
Yusuf DEMİR
Bakırçay University Çiğli Training and Research Hospital, Department of Cardiology, İzmir, Türkiye
yusufdemir2502@gmail.com
ORCID ID: 0000-0001-9167-493X

INTRODUCTION

Adults with an unrepaired secundum atrial septal defect (ASD) are predisposed to atrial arrhythmias due to chronic right atrial (RA) volume overload, progressive atrial dilation, and electrical remodeling, all of which contribute to the development of an arrhythmogenic substrate over time.[1-3] Although transcatheter ASD closure effectively eliminates the left-to-right shunt and reduces right-sided volume overload, early and late atrial tachyarrhythmias continue to occur following the procedure, particularly in older patients, those with delayed defect closure, and those with advanced atrial remodeling prior to the procedure.[4-6] Previous studies have suggested that persistent electrophysiological abnormalities may remain despite successful anatomical correction, indicating that eliminating the hemodynamic burden does not necessarily reverse established atrial remodeling.[7-9]

Recent evidence has further expanded the concept of atrial remodeling to encompass the broader entity of atrial cardiomyopathy, which includes structural, electrical, contractile, and autonomic abnormalities that predispose patients to atrial fibrillation (AF), thromboembolic events, and adverse

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

cardiovascular outcomes, even before clinically overt arrhythmia develops.[10] Consequently, identifying simple, reproducible, and widely available markers that reflect the underlying atrial substrate has become an important clinical objective. Surface electrocardiography (ECG) is an inexpensive and universally accessible tool that provides valuable information on atrial conduction abnormalities without requiring advanced imaging techniques.[11,12]

Among ECG-derived parameters, P-wave dispersion (PWDis)-defined as the difference between the maximum and minimum P-wave durations across the 12-lead ECG-reflects heterogeneous atrial conduction and has consistently been associated with paroxysmal AF, atrial conduction delay, and recurrent arrhythmias in various cardiovascular conditions.[13-17] Likewise, P-terminal force in lead V1 (PTFV1), originally described by Morris and colleagues, quantifies the duration and amplitude of the terminal negative portion of the P wave in that lead and has emerged as a robust electrocardiographic marker of left atrial electrical and structural remodeling. A growing body of evidence has demonstrated significant associations between abnormal PTFV1 and incident AF, ischemic stroke, atrial cardiomyopathy, and cardiovascular mortality across diverse patient populations.[18-21]

Despite the growing body of evidence supporting ECG-derived atrial biomarkers, data regarding the predictive value of PWDis and PTFV1 specifically in patients undergoing transcatheter ASD closure remain limited. Because early post-procedural AF may prolong hospitalization, increase the risk of thromboembolism, and necessitate additional rhythm management, identifying patients at increased risk prior to the procedure may facilitate individualized peri-procedural monitoring and preventive strategies. Therefore, the present study aimed to evaluate whether baseline PWDis and PTFV1 independently predict early (≤48-hour) AF following successful percutaneous secundum ASD closure.

METHODS

Ethics committee approval for the study was obtained from the Non-Interventional Clinical Research Ethics Committee of İzmir Bakırçay University (decision number: 2105; date: 7 March 2025). Written informed consent was not obtained because of retrospective design.

We conducted a retrospective cohort study at two tertiary centers. We included 555 consecutive adults (18-85 years) who underwent successful percutaneous secundum ASD closure between March 2015 and September 2024 at two cardiology clinics. Exclusion criteria were prior AF, any failed procedure, need for urgent cardiac surgery, known coronary artery disease, death within 48 hours of the procedure, suboptimal records, or unavailable preprocedural ECGs. Demographics, baseline 12-lead ECGs, transthoracic echocardiography, laboratory values, procedural details, and occluder device characteristics were extracted from hospital records. During hospitalization, all ECGs were acquired in the supine position at 25 mm/s and 10 mm/mV with 0.05-100 Hz filter settings, scanned and digitized, and analyzed at 4× magnification on a high-resolution monitor. All patients were continuously monitored with in-hospital telemetry during the first 24 hours following the procedure. During the subsequent 24 hours, patients were evaluated with standard 12-lead ECG recordings in the presence of symptoms or rhythm irregularities detected by bedside monitoring systems or wearable devices. AF was defined as an irregularly irregular rhythm without discernible P waves lasting ≥30 seconds on surface ECG. PWDis was calculated as the difference between the maximum and minimum P-wave durations across all 12 leads. PTFV1 was defined as the product of the duration and amplitude of the terminal negative P-wave component in lead V1 (μV·ms) (Figure 1). AF within the first 48 hours post-procedure was diagnosed by surface ECG or Holter monitoring. Patients without AF were classified as Group 1, and those with AF were classified as Group 2. ECG measurements were performed by two independent cardiologists who were blinded to clinical outcomes.

Statistical Analysis

Normality of continuous variables was assessed using the Kolmogorov-Smirnov test. Continuous variables are reported as mean ± standard deviation and categorical variables as counts and percentages. Group comparisons were performed using Student's t-test or Mann-Whitney U test for continuous variables, and chi-square or Fisher's exact test for categorical variables. Logistic regression identified independent ECG predictors of AF. Optimal cut-offs were derived by receiver-operating characteristic (ROC) analysis. Statistical significance was set at $p<0.05$.

RESULTS

Of 555 patients, 175 (31.5%) were male and 380 (68.5%) female; mean age was 39.1±15.3 years (range 18-85). At baseline, 510 (91.9%) patients were NYHA class I and 45 (8.1%) were class II. All procedures were technically successful; however, 27 patients (4.9%) developed AF within 48 hours. Patients with AF were



Figure 1. P-wave terminal force measurement in lead V1(PTFV1)=(a)P'duration×(b)P'amplitude

PTFV1: P wave terminal force in lead V1

significantly older (59.6±13.1 vs. 37.9±14.7 years, p<0.001), and AF was more common in NYHA class II (p<0.001). During the study period, no patients with NYHA class III-IV were deemed suitable for transcatheter closure. Mean ASD diameter was 15.8±5.7 mm; mean occluder size 21.5±6.2 mm. Mean LVEF was 61.7±3.2%, mean sPAP 42.4±16.6 mmHg, and mean Qp/Qs 2.2±0.6. There were no between-group differences in echocardiographic or laboratory parameters (all p>0.05) (Tables 1 and 2). Mean heart rate was 70.8±9.2 bpm; mean PWDis 36.7±11.9 ms; mean PTFV1 3904±666 μV·ms. Compared with Group 1, Group 2 had greater PWDis (42.8±13.9 vs. 36.4±11.8 ms, p=0.007) and higher PTFV1 (4325.9±730.4 vs. 3876.5±656.1 μV·ms, p=0.001) (Table 1). In multivariable models, both PWDis [odds ratio (OR) 1.036; 95% confidence interval (CI) 1.004-1.068; p=0.025] and PTFV1 (OR 1.086; 95% CI 1.027-1.149; p=0.004) independently predicted early AF (Table 3). ROC analysis yielded cut-offs for PTFV1 (>4,050 μV·ms; AUC 0.681; sensitivity 63%; specificity 62%) and for PWDis (>36.5 ms; AUC 0.634; sensitivity 57%; specificity 61%) (Figure 2). ROC analysis showed modest discriminatory performance.

Table 1. Baseline clinical, electrocardiographic and echocardiographic characteristics

Variables (mean ± SD)		**Group 1 n=528**	**Group 2 n=27**	**p value**
Age, years (mean ± SD)		37.9±14.7	59.6±13.1	**<0.001**
Male gender, n (%)		164 (31.1)	11 (40.7)	0.291
HT, n (%)		13 (2.5)	2 (7.4)	0.162
DM, n (%)		6 (1.1)	1 (3.7)	0.296
NYHA class	1	494 (93.6)	16 (59.2)	**<0.001**
	2	34 (6.4)	11 (40.8)	
Echocardiographic parameters				
ASD diameter (mm)		15.8±5.5	16.1±8.8	0.759
Device size (mm)		21.4±6.1	23.1±8.2	0.185
LVEDD (mm)		41.9±7.2	44.2±8.1	0.115
LVESD (mm)		28.9±5.9	28.8±6.4	0.897
LA (mm)		36.5±5.1	34.6±2.9	0.233
RA (mm)		32.8±8.6	31.5±3.5	0.619
LVEF (%)		61.7±3.2	61.9±3.1	0.806
RVEDD (mm)		35.6±9.1	36.6±10.5	0.653
SPAB (mmHg)		42.4±16.8	43.9±13.9	0.666
RV/LV ratio		0.85±0.2	0.91±0.3	0.336
RA/LA ratio		0.91±0.2	0.92±0.4	0.924
Qp/Qs ratio		2.2±0.6	2.3±0.4	0.609
Electrocardiographic parameters				
Heart rate (bpm)		71.2±9.3	68.2±8.4	0.309
Maximum P wave duration (ms)		101.5±10.2	105.1±10.7	0.076
Minimum P wave duration (ms)		65.1±7.4	62.2±8.8	0.056
P wave dispersion (ms)		36.4±11.8	42.8±13.9	**0.007**
PTFV1 (μV·ms)		3876.5±656.1	4325.9±730.4	**0.001**

SD: Standard deviation, n: Number of patients, ASD: Atrial septal defect, HT: Hypertension, LA: Left atrium, LV: Left ventricle, LVEDD: Left ventricular end-diastolic diameter, LVEF: Left ventricular ejection fraction, LVESD: Left ventricular end-systolic diameter, DM: Diabetes mellitus, NYHA: New York Heart Association classification, PTFV1: P wave terminal force in lead V1, RA: Right atrium, RV: Right ventricle, RVEDD: Right ventricular end-diastolic diameter, SPAB: Systolic pulmonary artery pressure

Table 2. Laboratory characteristics of patients according to groups			
Variables (mean ± SD)	**Group 1 n=528**	**Group 2 n=27**	**p value**
Urea, mg/dL	27.5±10.8	30.3±6.4	0.426
Creatinine, mg/dL	0.7±0.2	0.65±0.1	0.212
Sodium, mEq/L	139.4±2.5	139.4±2.9	0.951
Potassium, mEq/L	4.3±0.4	4.2±0.4	0.686
Calcium, mg/dL	9.2±0.4	8.9±0.6	0.088
WBC, $\times 10^9$/L	7.76±2.1	8.21±1.1	0.685
Neutrophil count, $\times 10^9$/L	4.75±1.8	4.9±0.8	0.879
Lymphocyte count, $\times 10^9$/L	2.14±0.7	2.37±0.4	0.506
Hemoglobin, g/dL	13.3±1.7	13.6±1.9	0.419
Hematocrit, %	40.1±5.5	40.5±7.1	0.777
Platelet count, $\times 10^9$/L	239.18±89.2	282.55±48.4	0.159
SD: Standard deviation, n: Number of patients, WBC: White blood cell count			

Table 3. Effects of ECG indicators on the development of atrial fibrillation in univariate and multivariate logistic regression analysis				
Variables	**Univariate logistic regression**		**Multivariate logistic regression**	
	OR (95% CI)	**p value**	**OR (95% CI)**	**p value**
Maximum P wave duration	1.034 (0.996-1.074)	0.078		
Minimum P wave duration	0.951 (0.904-1.001)	0.056		
P wave dispersion	1.042 (1.011-1.074)	**0.008**	1.036 (1.004-1.068)	**0.025**
PTFV1	1.096 (1.038-1.157)	**0.001**	1.086 (1.027-1.149)	**0.004**
PTFV1: P wave terminal force in lead V1, CI: Confidence interval, OR: Odds ratio				

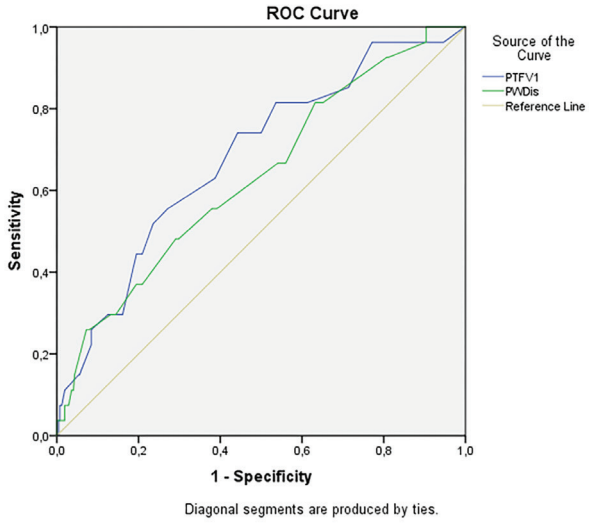


Figure 2. Receiver operating characteristic (ROC) curve analysis of the optimal cut-off values of P-wave dispersion and P-wave terminal force in lead V1

PWDis: P-wave dispersion, PTFV1: P wave terminal force in lead V1

DISCUSSION

In this two-center cohort of adults undergoing transcatheter secundum ASD closure, we demonstrated that two readily obtainable ECG-derived parameters, PWDis and PTFV1, independently predicted early AF within 48 hours after the procedure. Although the incidence of early AF was relatively low, patients who developed AF exhibited significantly more pronounced abnormalities in both ECG markers measured before intervention, suggesting that pre-existing atrial electrical remodeling, rather than the procedure itself, is the principal determinant of early arrhythmogenesis. These findings support the growing concept that conventional surface ECG may provide clinically meaningful information regarding atrial substrate before structural heart interventions.

The biological plausibility of our findings is supported by the well-established pathophysiology of ASD. Chronic left-to-right shunting exposes the right atrium to prolonged volume overload, resulting in atrial dilatation, myocyte stretch, neurohormonal activation, inflammation, and progressive fibrosis. These changes alter conduction velocity and increase conduction heterogeneity, facilitating re-entry and AF. Importantly, several experimental and clinical studies have demonstrated that electrical remodeling may persist despite successful defect closure, particularly in adults undergoing late intervention.[22-26] Consequently, elimination

of the hemodynamic burden does not necessarily normalize the pre-existing arrhythmogenic substrate.

PWDis reflects heterogeneous atrial conduction and has consistently been associated with AF in several cardiovascular disorders.[27-31] Increased dispersion indicates delayed and non-uniform atrial activation, thereby facilitating wavefront fragmentation and re-entry. Our results extend previous observations by demonstrating that PWDis remains clinically relevant in patients undergoing transcatheter ASD closure, indicating that atrial conduction abnormalities persist despite correction of the underlying congenital lesion.

PTFV1 has emerged as one of the most robust ECG markers of atrial cardiomyopathy. Unlike simple P-wave duration, PTFV1 reflects both delayed atrial activation and increased left atrial electrical load, and correlates with atrial fibrosis, AF, stroke, and cardiovascular outcomes.[32-35] Although ASD primarily affects the right heart, long-standing interatrial interactions and biatrial remodeling may explain why an index reflecting left atrial abnormalities independently predicted early AF in our cohort. These findings further support the concept that atrial electrical remodeling in ASD represents a biatrial, rather than an exclusively RA, process.

The incidence of early AF in our cohort (4.9%) is comparable to that reported in previous studies evaluating transcatheter ASD closure, in which post-procedural AF rates have generally ranged between approximately 2% and 7%, depending on patient age, timing of rhythm assessment, and duration of follow-up. Older age has consistently emerged as one of the strongest determinants of post-closure atrial arrhythmias, supporting the concept that the burden of pre-existing atrial remodeling, rather than the procedure itself, primarily drives arrhythmia occurrence.[4,5,22,23] The present findings extend these observations by demonstrating that ECG-derived markers of atrial electrical remodeling provide additional prognostic information beyond age alone.

The modest ROC performance observed in our analysis should not diminish the clinical value of these markers. AF is a multifactorial disease, and excellent discrimination based on a single ECG variable is not expected. Instead, PWDis and PTFV1 should be regarded as complementary biomarkers that may improve prediction when integrated with age, atrial dimensions, clinical variables and procedural characteristics.

Another noteworthy finding of the present study is that neither echocardiographic parameters nor routine laboratory parameters differed significantly between patients who developed early AF and those who remained in sinus rhythm, whereas both ECG-derived indices retained independent predictive value. This observation suggests that subtle electrical remodeling may precede overt structural abnormalities detectable by conventional imaging. Similar findings have been reported in previous studies, which demonstrate that electrocardiographic markers often identify atrial vulnerability before measurable changes in atrial size or function become evident.[10,18,20] Accordingly, PWDis and PTFV1 should not be regarded merely as surrogate ECG measurements, but rather as noninvasive indicators of the underlying atrial substrate. This concept is particularly relevant in adults with ASD, in whom chronic atrial remodeling develops gradually over many years and may persist despite successful defect closure. Therefore, the incorporation of ECG-derived atrial biomarkers into routine preprocedural assessment may complement conventional echocardiographic evaluation and improve the identification of patients at increased risk of early post-procedural AF.

From a practical perspective, both PWDis and PTFV1 are inexpensive, universally available, and require only a standard 12-lead ECG. Identification of high-risk patients before ASD closure may justify closer telemetry surveillance, aggressive correction of reversible peri-procedural triggers, and consideration of prolonged ambulatory rhythm monitoring after discharge. Such individualized management may facilitate earlier AF detection and timely initiation of appropriate therapy.

Although early AF may be clinically relevant in the peri-procedural setting, it likely reflects a combination of pre-existing atrial substrate and transient triggers such as procedural inflammation, autonomic imbalance, and acute hemodynamic shifts. Therefore, our results do not necessarily translate to medium- or long-term arrhythmia burden.

Study Limitations

Some limitations of this study should be acknowledged. First, the retrospective design has limitations. Second, rhythm monitoring was performed continuously via telemetry during the first 24 hours after the procedure, while AF detection during the subsequent 24-hour period relied on bedside monitoring systems and symptom-triggered ECG recordings. Although this monitoring strategy reflects routine clinical practice in many institutions, it may have failed to detect short-term or asymptomatic AF episodes. Third, PWDis and PTFV1 measurements were obtained from digitized ECG recordings by manually defining the onset and end of the P wave. Although all measurements were performed at high magnification by two independent cardiologists blinded to clinical outcomes, manual ECG analysis may be subject to measurement variability. Finally, the relatively small number of AF events limited the complexity of multivariate modeling and may have increased the risk of model instability or overfitting. Therefore, the number of predictors included in the regression model was intentionally limited. Accordingly, the findings should be interpreted with caution and primarily considered hypothesis-generating. Furthermore, ROC analyses demonstrated only modest discriminatory performance, suggesting that these ECG indices alone are insufficient for clinical decision-making but may contribute as complementary markers within a broader risk assessment framework. Taken

together, our findings should be considered hypothesis-generating and should be confirmed in larger prospective studies with extended rhythm follow-up.

The present findings also emphasize that ECG-derived atrial biomarkers should be interpreted as markers of the underlying atrial substrate rather than isolated predictors of arrhythmia. Although the predictive performance of PWDis and PTFV1 was modest, their widespread availability, low cost, and ease of measurement make them attractive adjunctive tools for identifying patients who may benefit from intensified rhythm surveillance after transcatheter ASD closure.

CONCLUSION

Baseline PWDis and PTFV1 independently predict early AF after percutaneous ASD closure. These readily obtainable ECG indices may enhance risk stratification and inform peri-procedural management in contemporary practice.

Ethics

Ethics Committee Approval: Ethics committee approval for the study was obtained from the Non-Interventional Clinical Research Ethics Committee of İzmir Bakırçay University (decision number: 2105; date: 7 March 2025).

Informed Consent: Written informed consent was not obtained because of retrospective design.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Y.D., S.S., O.Ş., Concept: A.A.B., Design: Y.D., O.Ş., Data Collection or Processing: Y.D., A.A.B., S.S., Analysis or Interpretation: O.Ş., Literature Search: A.A.B., Writing: A.A.B.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Gatzoulis MA, Freeman MA, Siu SC, Webb GD, Harris L. Atrial arrhythmia after surgical closure of atrial septal defects in adults. N Engl J Med. 1999;340:839-46.
2. Berger F, Vogel M, Kramer A, et al. Incidence of atrial flutter/fibrillation in adults with atrial septal defect before and after surgery. Ann Thorac Surg. 1999;68:75-8.
3. Oliver JM, Gallego P, González A, Benito F, Mesa JM, Sobrino JA. Predisposing conditions for atrial fibrillation in atrial septal defect with and without operative closure. Am J Cardiol. 2002;89:39-43.
4. Giardini A, Donti A, Sciarra F, Bronzetti G, Mariucci E, Picchio FM. Long-term incidence of atrial fibrillation and flutter after transcatheter atrial septal defect closure in adults. Int J Cardiol. 2009;134:47-51.
5. Spies C, Khandelwal A, Timmermanns I, Schräder R. Incidence of atrial fibrillation following transcatheter closure of atrial septal defects in adults. Am J Cardiol. 2008;102:902-6.
6. Abaci A, Unlu S, Alsancak Y, Kaya U, Sezenoz B. Short and long term complications of device closure of atrial septal defect and patent foramen ovale: meta-analysis of 28,142 patients from 203 studies. Catheter Cardiovasc Interv. 2013;82:1123-38.
7. Morton JB, Sanders P, Vohra JK, et al. Effect of chronic right atrial stretch on atrial electrical remodeling in patients with an atrial septal defect. Circulation. 2003;107:1775-82.
8. Fang F, Luo XX, Lin QS, et al. Characterization of mid-term atrial geometrical and electrical remodeling following device closure of atrial septal defects in adults. Int J Cardiol. 2013;168:467-71.
9. Chubb H, Whitaker J, Williams SE, et al. Pathophysiology and management of arrhythmias associated with atrial septal defect and patent foramen ovale. Arrhythm Electrophysiol Rev. 2014;3:168-72.
10. Goette A, Corradi D, Dobrev D, et al. Atrial cardiomyopathy revisited-evolution of a concept: a clinical consensus statement of the European Heart Rhythm Association (EHRA) of the ESC, the Heart Rhythm Society (HRS), the Asian Pacific Heart Rhythm Society (APHRS), and the Latin American Heart Rhythm Society (LAHRS). Europace. 2024;26:euae204.
11. Baumgartner H, De Backer J, Babu-Narayan SV, et al. 2020 ESC Guidelines for the management of adult congenital heart disease. Eur Heart J. 2021;42:563-645.
12. Johnson JN, Marquardt ML, Ackerman MJ, et al. Electrocardiographic changes and arrhythmias following percutaneous atrial septal defect and patent foramen ovale device closure. Catheter Cardiovasc Interv. 2011;78:254-61.
13. Dilaveris PE, Gialafos JE. P-wave dispersion: a novel predictor of paroxysmal atrial fibrillation. Ann Noninvasive Electrocardiol. 2001;6:159-65.
14. Aytemir K, Ozer N, Atalar E, et al. P wave dispersion on 12-lead electrocardiography in patients with paroxysmal atrial fibrillation. Pacing Clin Electrophysiol. 2000;23:1109-12.
15. Ozer N, Aytemir K, Atalar E, et al. P wave dispersion in hypertensive patients with paroxysmal atrial fibrillation. Pacing Clin Electrophysiol. 2000;23:1859-62.
16. Okutucu S, Aytemir K, Oto A. P-wave dispersion: What we know till now? JRSM Cardiovasc Dis. 2016;5:2048004016639443.
17. Magnani JW, Williamson MA, Ellinor PT, Monahan KM, Benjamin EJ. P wave indices: current status and future directions in epidemiology, clinical, and research applications. Circ Arrhythm Electrophysiol. 2009;2:72-9.
18. Chen LY, Ribeiro ALP, Platonov PG, et al. P wave parameters and indices: a critical appraisal of clinical utility, challenges, and future research-a consensus document endorsed by the International Society of Electrocardiology and the International Society for Holter and Noninvasive Electrocardiology. Circ Arrhythm Electrophysiol. 2022;15:e010435.
19. Bethell HJ, Nixon PG. P wave of electrocardiogram in early ischaemic heart disease. Br Heart J. 1972;34:1170-5.
20. Wolder LD, Graff C, Baadsgaard KH, et al. Electrocardiographic P terminal force in lead V1, its components, and the association with stroke and atrial fibrillation or flutter. Heart Rhythm. 2023;20:354-62.
21. Eranti A, Aro AL, Kerola T, et al. Prevalence and prognostic significance of abnormal P terminal force in lead V1 of the ECG in the general population. Circ Arrhythm Electrophysiol. 2014;7:1116-21.
22. Woo SB, Jang SI, Kim SH, et al. Transcatheter closure of secundum atrial septal defect in patients over 60 years old. Korean Circ J. 2013;43:110-4.
23. Wagdi P. Incidence and predictors of atrial fibrillation following transcatheter closure of interatrial septal communications using contemporary devices. Clin Res Cardiol. 2010;99:507-10.

24. Bronzetti G, D'Angelo C, Donti A, et al. Role of atrial fibrillation after transcatheter closure of patent foramen ovale in patients with or without cryptogenic stroke. Int J Cardiol. 2011;146:17-21.
25. Berger F, Vogel M, Kretschmar O, Dave H, Prêtre R, Dodge-Khatami A. Arrhythmias in patients with surgically treated atrial septal defects. Swiss Med Wkly. 2005;135:175-8.
26. Deaconu S, Deaconu A, Marascu G, et al. Arrhythmic risk and treatment after transcatheter atrial septal defect closure. Diagnostics (Basel). 2023;14:33.
27. Doğduş M, Turan OE, Başkurt AA, et al. An effective novel index for predicting the recurrence of atrial fibrillation ablation: p wave duration-to-amplitude ratio. Turk Kardiyol Dern Ars. 2022;50:498-504.
28. Antoun I, Li X, Vali Z, et al. The value of P-wave parameters changes in predicting catheter ablation outcomes for paroxysmal atrial fibrillation. Ann Noninvasive Electrocardiol. 2025;30:e70047.
29. Masood S, Ashraf SMK, Malik MA, Wahab S. P-wave indices and left atrial mechanics as predictors of atrial cardiopathy in embolic stroke of undetermined source. Sci Rep. 2023;13:19965.
30. Intzes S, Zagoridis K, Symeonidou M, et al. P-wave duration and atrial fibrillation recurrence after catheter ablation: a systematic review and meta-analysis. Europace. 2023;25:450-9.
31. Maezono A, Oishi E, Sasaki T, et al. Electrocardiographic P-wave indices and onset of atrial fibrillation and histological fibrosis in a general Japanese population: a prospective cohort analysis of the Hisayama Study. Eur Heart J Open. 2025;5:oeaf126.
32. Arslan GY, Söner S. P-wave terminal force in lead V1-Morris index after catheter ablation: a study in atrial fibrillation and typical atrial flutter patients. Eskisehir Med J. 2026;7:134-7.
33. Sudo Y, Morimoto T, Tsushima R, et al. P-wave terminal force in lead V1 and outcomes in patients with persistent atrial fibrillation undergoing catheter ablation. Am Heart J. 2023;141-150.
34. Wang Z, Wang B, Yang Y, Yang X, Che Y, Xia Y. P-wave terminal force in lead V1 is associated with recurrence after catheter ablation in patients with paroxysmal atrial fibrillation and normal left atrial size. Front Cardiovasc Med. 2024;11:1467585.
35. Tereshchenko LG. P-terminal force in lead V1: May the force be with you. Heart Rhythm. 2023;20:363-4.

DOI: 10.4274/forbes.galenos.2026.84758
Forbes J Med 2026;7:178-184

Original Article

Fetal Thymic-to-Thoracic Ratio and Neonatal Adaptation in Pregnancies Complicated by Gestational Diabetes Mellitus: A Prospective Cohort Study

Raziye TORUN[1], Emine DEMİR[2], Sevim TUNCER CAN[3], Zübeyde EMİRALİOĞLU ÇAKIR[1]

[1]University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Perinatology, İzmir, Türkiye
[2]University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital, Clinic of Perinatology, İzmir, Türkiye
[3]University of Health Sciences Türkiye, Gazi Yaşargil Training and Research Hospital, Clinic of Perinatology, Diyarbakır, Türkiye

Cite as: Torun R, Demir E, Tuncer Can S, Emiralioğlu Çakır Z. Fetal thymic-to-thoracic ratio and neonatal adaptation in pregnancies complicated by gestational diabetes mellitus: a prospective cohort study. Forbes J Med. 2026;7:178-184

ABSTRACT

Objective: To compare fetal thymus biometry among pregnant women with insulin-requiring gestational diabetes mellitus (GDM), diet-regulated GDM, and healthy controls, and to investigate the relationship between the fetal thymic-thoracic ratio (TTR) and 1st and 5th minute Apgar scores.

Methods: This prospective cohort study included 145 singleton pregnancies between 28 and 36 weeks of gestation: 44 with insulin-requiring GDM, 46 with diet-regulated disease, and 55 healthy controls with a normal 75 g oral glucose tolerance test. Fetal thymic anteroposterior (AP) and transverse diameters (TTD) and the thoracic AP diameter were measured sonographically, and TTR was calculated. Groups were compared with the Kruskal-Wallis test and associations by Spearman rank correlation.

Results: The groups did not differ in AP thymic diameter (p=0.371), TTD (p=0.844) or TTR (p=0.638). In the whole cohort, the TTR correlated weakly and positively with Apgar scores at the 1st (ρ=0.226, p=0.006) and 5th minutes (ρ=0.241, p=0.004). Correlations were moderate in the insulin-requiring group (ρ=0.410, p=0.006 and ρ=0.409, p=0.006, respectively), although pairwise comparison of coefficients showed no significant difference between groups. The AP thymic diameter correlated weakly with both scores, whereas the TTD showed no significant correlation.

Conclusion: Fetal thymus dimensions and TTR were found to be similar among the three study groups. The association of higher TTR values with higher 1st and 5th minute Apgar scores suggests that fetal thymus measurements normalized to thoracic dimensions may be sonographic parameter associated with early neonatal adaptation.

Keywords: Gestational diabetes mellitus, fetal thymus, thymic-to-thoracic ratio, Apgar score, fetal ultrasonography

Received: 11.09.2026
Accepted: 14.10.2026
Publication Date: 19.10.2026

Corresponding Author:
Raziye TORUN
University of Health Sciences Türkiye, İzmir City Hospital, Clinic of Perinatology, İzmir, Türkiye
dr.razzz_03@hotmail.com
ORCID ID: 0000-0002-0272-7196

INTRODUCTION

Gestational diabetes mellitus (GDM) characterized by carbohydrate intolerance that is first detected or that manifests during pregnancy ranks among the most prevalent metabolic disorders encountered in obstetric practice.[1] Maternal hyperglycemia may precipitate adverse maternal-fetal outcomes, including fetal hyperinsulinemia, accelerated somatic growth, macrosomia, intrauterine hypoxia, and birth trauma.[2]

First-line therapy for GDM centers on dietary adjustment and lifestyle modification (diet-controlled GDM); however, patients who fail to meet glycemic targets require insulin therapy (insulin-requiring GDM). Pregnancies requiring insulin are typically characterized by severe peripheral insulin resistance, a more pronounced systemic inflammatory response, and vascular and metabolic alterations that may impair fetal tissue perfusion.[3]

The fetal thymus serves as the primary lymphoid organ of the immune system and can be visualized via transabdominal ultrasound on the three-vessel view from the second trimester onward. A key characteristic of the thymus is its rapid morphological regression (involution) in response to acute or chronic stress. Accordingly, a reduced thymic size has been proposed as an indirect marker of fetal inflammatory response syndrome and exposure to adverse intrauterine stress.[4]

Because absolute thymic diameters correlate with advancing gestational age, the thymic-thoracic ratio (TTR), defined as the ratio of the anteroposterior (AP) thymic diameter to the

Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Buca Seyfi Demirsoy Training and Research Hospital.
This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

intrathoracic mediastinal diameter, was introduced to quantify thymic alterations independently of gestational age. TTR offers the distinct advantage of standardizing relative fetal thymic volume along a uniform axis, unaffected by gestational age or generalized fetal growth trajectories. Although thymic involution has been well documented in adverse obstetric conditions including preterm premature membrane rupture, hypertensive disorders of pregnancy, and fetal growth impairment, the direction of thymic remodeling in GDM specifically across diet-controlled and insulin-requiring subgroups and how these morphological changes relate to immediate neonatal vigor (Apgar scores) remain ill-defined.[5,6]

The Apgar score is the most fundamental clinical indicator of a newborn's postnatal physiological adaptation and cardiorespiratory stability. Several prospective cohort investigations have documented an elevated likelihood of depressed Apgar scores at both the 1st and 5th minutes of life among neonates born to women with GDM.[7] However, no practical marker currently exists that can prenatally predict which GDM-complicated pregnancies will result in adverse neonatal outcomes.

This study was designed to determine whether a relationship exists between fetal TTR and neonatal Apgar scores (at 1 and 5 minutes) among healthy pregnant women and women with diet-controlled or insulin-requiring GDM. Evaluate the predictive utility of absolute fetal thymic diameters (AP and transverse) for neonatal Apgar outcomes; We compare fetal thymic biometry across modalities of glycemic management.

METHODS

This prospective cohort study was conducted between June 2026 and August 2026 at a tertiary perinatology center. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and approved by the Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital (approval number: 2026/05-19; date: 04.06.2026). Prior to participation in the study, all participants were provided with detailed information, and written informed consent was obtained from each participant.

Singleton pregnancies between 28 and 36 completed weeks' gestation were recruited. Three study arms were defined: a diet-controlled GDM cohort (n=46), an insulin-requiring GDM cohort (n=44), and a gestational-age-matched healthy control group (n=55), whose normoglycemia was confirmed by a 75-g oral glucose tolerance test (OGTT) conducted at 24-28 weeks' gestation. To ensure an uncomplicated baseline profile, women with a history of obstetric complications in previous pregnancies, including preterm delivery, gestational hypertension, intrauterine growth restriction, and preeclampsia, were excluded from the control group. Exclusion criteria included maternal systemic disease, multiple pregnancies, fetal structural or genetic anomalies, fetal growth restriction, and maternal smoking. Women who received antenatal corticosteroids to promote fetal lung maturation were also excluded because of the established effect of glucocorticoids on fetal thymic involution.

Power calculations were carried out with G*Power 3.1, targeting the primary outcome of the TTR-Apgar correlation. With an alpha level of 0.05, 80% power, and an expected effect size of r=0.235, at least 138 participants were required. To account for potential loss to follow-up, 145 singleton pregnancies were enrolled (55 controls, 46 diet-controlled GDM, and 44 insulin-requiring GDM). This final sample size provided 80% power (α=0.05) to detect a moderate effect size (f=0.26) in three-group comparisons.

Universal GDM screening was performed using a one-step 75-g OGTT at 24-28 weeks' gestation. Gestational age was confirmed by combining crown-rump length measurement during first-trimester ultrasonography with dating from the last menstrual period. Women meeting one or more plasma glucose thresholds fasting ≥92 mg/dL, 1-hour post-load ≥180 mg/dL, or 2-hour post-load ≥153 mg/dL were classified as having GDM in accordance with International Association of Diabetes and Pregnancy Study Groups criteria.[8]

Sonographic scans and measurements were conducted by a single seasoned maternal-fetal medicine specialist to preclude inter-observer discrepancy, utilizing a real-time Samsung HS70A ultrasound device (Samsung Medison, Seoul, South Korea; 2017).

On the axial three-vessel view, the fetal thymus appeared as a uniformly echogenic structure in the anterior mediastinum. The thymic transverse diameter (TTD) was obtained at the greatest transverse dimension perpendicular to an imaginary line drawn from the sternal midpoint to the vertebral body.[9] For the TTR calculation, the AP thymic diameter was defined as the distance from the inner sternal cortex to the leading edge of the transverse aortic arch, and the intrathoracic (AP thoracic) diameter was defined as the distance between the inner sternal cortex and the anterior vertebral margin in the same plane (Figure 1). TTR was computed as the ratio of the thymic AP diameter to the total thoracic AP diameter.

Statistical Analysis

All statistical procedures were executed in IBM SPSS Statistics (version 31.0; IBM Corp., Armonk, NY, USA). The normality of continuous variables was tested with the Shapiro-Wilk method; non-normally distributed data were reported as the median and interquartile range (25th-75th percentile), whereas categorical data were presented as counts and percentages [n (%)]. Non-parametric group comparisons were made with the Kruskal-Wallis test, and where a global difference was detected, pairwise post-hoc analyses incorporating Bonferroni correction for multiplicity were performed. Categorical variables were analyzed using Pearson's chi-square test or Fisher's exact test, depending on expected cell frequencies. Spearman rank correlation coefficients were computed to quantify the strength of the TTR-Apgar relationship, initially across the full cohort and subsequently within each

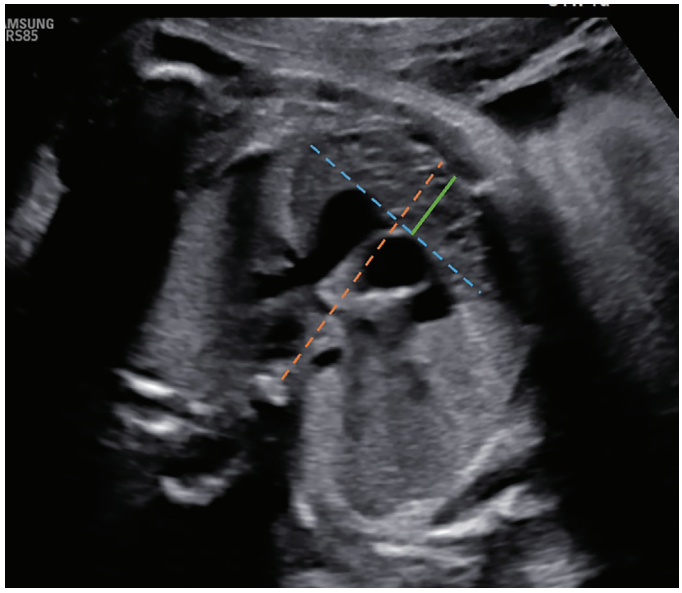


Figure 1. Measurement of the fetal thymus and thorax on the axial three-vessel view (two-dimensional grayscale ultrasonography). Green solid line: thymic anteroposterior diameter; orange dashed line: thoracic anteroposterior diameter; blue dashed line: thymic transverse diameter

subgroup. The magnitude of each correlation coefficient was classified as negligible (0.00-0.10), weak (0.10-0.39), moderate (0.40-0.69), strong (0.70-0.89), or very strong (0.90-1.00) following Schober et al.[10] To test for statistically significant divergence among subgroup-specific correlation values, Fisher's r-to-z transformation was applied.

All statistical tests were two-sided and a p value <0.05 was considered statistically significant. Adjusted p values were reported for post-hoc multiple comparisons.

RESULTS

The final study population comprised 145 pregnant women: 44 (30.3%) with insulin-requiring GDM, 46 (31.7%) with diet-controlled GDM, and 55 (37.9%) serving as healthy controls.

The demographic and obstetric features of all three cohorts are presented in Table 1. The median gestational age at which sonographic thymus assessment was performed was 32.0 weeks across all groups, and measurements were carried out at similar gestational ages. Maternal age differed significantly among the cohorts (p=0.031), and post-hoc pairwise comparisons revealed that this disparity was primarily attributable to the difference between the insulin and control groups (adjusted p=0.042). Maternal body mass index (BMI) also varied significantly across groups (p<0.001), being lower in the control group than in both the insulin and diet cohorts. Gravidity, parity, and rates of prior pregnancy loss were comparable across all three arms.

Fetal biometric and thymic measurements are summarized in Table 2. No significant differences were observed among the three groups regarding AP thymic diameter (p=0.371) or TTD (p=0.844). Similarly, the TTR was comparable across cohorts, with values of 0.405 (0.336-0.442) in the insulin-requiring group, 0.407 (0.368-0.442) in the diet-controlled group, and 0.395 (0.352-0.441) in the control group; there was no statistically significant intergroup variation (p=0.638). BPD and HC were equivalent across groups. Although omnibus comparisons demonstrated statistical significance for AC (p=0.032) and FL (p=0.041), post-hoc pairwise comparisons revealed no individual contrast that remained significant after Bonferroni correction. Thoracic diameter differed significantly across cohorts (p=0.033), and post-hoc testing demonstrated larger thoracic diameters in the insulin-requiring group than in controls (adjusted p=0.039).

Delivery occurred at a significantly earlier gestational age in the insulin-requiring group compared with controls (adjusted p=0.002; overall p=0.002). Neonatal birth weight likewise differed among groups (p=0.008); infants born to insulin-treated mothers were heavier than those in the control arm (adjusted p=0.007). Neither the route of delivery (p=0.233) nor the fetal sex distribution (p=0.395) varied significantly across cohorts. Apgar scores at 1 and 5 minutes were similar across the three groups (p=0.588 and p=0.818, respectively) (Table 3).

Regarding the primary outcome, the relationship between TTR and Apgar scores was analyzed. Across the entire study population, a weak but statistically significant positive correlation was identified between TTR and both 1st-minute Apgar scores (ρ=0.226, p=0.006) and 5th-minute Apgar scores (ρ=0.241, p=0.004) (Table 4). When

Table 1. Comparison of maternal and obstetric characteristics between groups

Variable	Insulin-requiring (n=44)	Diet-controlled (n=46)	Control (n=55)	p
Maternal age, years	30.0 (28.0-34.2)	28.0 (25.2-31.0)	27.0 (24.0-33.0)	**0.031**
BMI, kg/m^2	31.9 (28.4-36.8)	30.6 (27.9-34.9)	28.3 (25.9-30.0)	**<0.001**
Gravidity	2.5 (2.0-4.0)	2.0 (1.0-3.0)	2.0 (1.0-3.0)	0.054
Parity	1.0 (1.0-2.0)	1.0 (0.0-2.0)	1.0 (0.0-1.5)	0.200
Previous abortions	0.0 (0.0-1.0)	0.0 (0.0-0.8)	0.0 (0.0-0.0)	0.186
Gestational age at scan, weeks	32.0 (30.8-34.0)	32.0 (30.0-34.0)	32.0 (30.0-33.0)	0.685

Data are presented as median (25th-75th percentile). Between-group comparisons were performed with the Kruskal-Wallis test. p values in bold indicate statistical significance (p<0.05). BMI: Body mass index

cohorts were evaluated individually, this association was most pronounced within the insulin-requiring GDM group, where TTR demonstrated moderate positive correlations with 1st- and 5th-minute Apgar scores (ρ=0.410, p=0.006 and ρ=0.409, p=0.006, respectively). In contrast, no significant correlations between TTR and Apgar scores were observed in the diet-controlled or control cohorts. However, direct intergroup comparisons of correlation coefficients yielded no statistically significant difference (p>0.14), precluding a definitive conclusion that this relationship is exclusive to the insulin-requiring cohort. For secondary outcomes, weak positive correlations were detected in the overall cohort between AP thymic diameter and both 1st-minute Apgar scores (ρ=0.182, p=0.028) and 5th-minute Apgar scores (ρ=0.201, p=0.015). Although these associations were also evident in the insulin group, they did not reach statistical significance in the diet or control groups. TTD showed no significant correlation with 1st- or 5th-minute Apgar scores either in the entire population or within separate cohorts.

Overall, although fetal TTR remained comparable across the three study groups, higher TTR values were weakly correlated with improved 1st- and 5th-minute Apgar scores. At the group level, this relationship was most evident in the insulin-requiring GDM

Table 2. Comparison of fetal biometry and thymus measurements between groups

Variable	Insulin-requiring (n=44)	Diet-controlled (n=46)	Control (n=55)	p
BPD, mm	84.0 (78.2-88.7)	81.0 (77.5-86.0)	82.7 (76.9-87.7)	0.511
HC, mm	304.1 (290.1-323.1)	292.0 (288.2-316.6)	300.6 (285.9-314.3)	0.227
AC, mm	292.2 (276.0-330.4)	279.5 (262.2-299.7)	283.8 (262.7-301.9)	**0.032**
FL, mm	63.1 (60.2-66.6)	58.5 (56.6-65.2)	59.9 (56.8-64.7)	**0.041**
Thymic AP diameter, mm	16.2 (13.6-18.7)	15.1 (13.4-17.8)	15.2 (12.4-16.9)	0.371
Thymic transverse diameter, mm	26.9 (23.6-30.9)	26.6 (23.6-30.4)	27.6 (24.4-30.2)	0.844
Thoracic diameter, mm	41.3 (36.5-47.7)	37.8 (33.8-41.6)	38.0 (33.0-42.9)	**0.033**
Thymic-thoracic ratio	0.405 (0.336-0.442)	0.407 (0.368-0.442)	0.395 (0.352-0.441)	0.638

Data are presented as median (25th-75th percentile). Between-group comparisons were performed with the Kruskal-Wallis test. p values in bold indicate statistical significance (p<0.05). BPD: Biparietal diameter, HC: Head circumference, AC: Abdominal circumference, FL: Femur length, AP: Anteroposterior

Table 3. Comparison of delivery and neonatal outcomes between groups

Variable	Insulin-requiring (n=44)	Diet-controlled (n=46)	Control (n=55)	p
Gestational age at delivery, weeks	37.0 (37.0-38.0)	38.0 (37.0-38.0)	38.0 (37.5-39.0)	**0.002**
Birth weight, g	3440 (3182-3750)	3305 (2681-3558)	3180 (2794-3520)	**0.008**
Cesarean delivery, n (%)	33 (75.0)	33 (71.7)	33 (60.0)	0.233
Male sex, n (%)	26 (59.1)	27 (58.7)	26 (47.3)	0.395
Apgar score at 1st minute	8.0 (8.0-8.0)	8.0 (8.0-8.0)	8.0 (7.5-8.0)	0.588
Apgar score at 5th minutes	9.0 (9.0-9.0)	9.0 (9.0-9.0)	9.0 (9.0-9.0)	0.818

Continuous variables are presented as median (25th-75th percentile) and categorical variables as n (%). Continuous variables were compared with the Kruskal-Wallis test and categorical variables with the chi-square test. p values in bold indicate statistical significance (p<0.05)

Table 4. Spearman correlations between thymus measurements and Apgar scores

Measurement	Apgar score at 1st minute				Apgar score at 5th minutes			
	All	Insulin-requiring	Diet-controlled	Control	All	Insulin-requiring	Diet-controlled	Control
Thymic-thoracic ratio	**0.226 (0.006)**	**0.410 (0.006)**	0.135 (0.371)	0.132 (0.336)	**0.241 (0.004)**	**0.409 (0.006)**	0.149 (0.324)	0.150 (0.275)
Thymic AP diameter	**0.182 (0.028)**	**0.329 (0.029)**	0.175 (0.246)	0.062 (0.653)	**0.201 (0.015)**	**0.316 (0.037)**	0.171 (0.256)	0.130 (0.345)
Thymic transverse diameter	0.092 (0.271)	0.129 (0.404)	0.101 (0.505)	0.054 (0.695)	0.132 (0.114)	0.111 (0.473)	0.094 (0.533)	0.172 (0.210)

Values are Spearman correlation coefficients (p value). Values in bold indicate statistical significance (p<0.05). No correction for multiple comparisons was applied; subgroup analyses are exploratory. AP: Anteroposterior

cohort. Thymic AP diameter displayed a weaker association with Apgar scores, whereas TTD showed no significant correlation with neonatal scores.

DISCUSSION

In this prospective investigation, we evaluated the course of fetal thymic biometry and TTR in groups with insulin-requiring GDM, diet-controlled GDM, and healthy controls, as well as the relationship between these measurements and neonatal Apgar scores. The main finding of our study was that TTR, which did not differ significantly between groups, was positively associated with both the 1st- and 5th-minute Apgar scores. In the subgroup analysis, this positive trend was most pronounced in cases of insulin-requiring GDM. However, the inability to demonstrate a significant difference in the direct comparison of correlation coefficients between groups precludes interpretation of this finding as specific to pregnant women using insulin. Given that a weaker association was observed between the AP diameter of the thymus and Apgar scores, and no significant association was found between the TTD and Apgar scores, this suggests that TTR, which normalizes fetal thymus size relative to thoracic dimensions, carries a different clinical significance than isolated millimeter diameters.

Available literature regarding the impact of maternal diabetes on fetal thymic development remains inconsistent; while several cohorts describe thymic enlargement, the prevailing evidence indicates that fetal thymic dimensions are diminished in diabetic gestations.[11,12]

Dörnemann et al.[13] reported that TTR levels were reduced in diabetic pregnant women but that there was no statistically significant difference among the three subgroups (diet-controlled GDM, insulin-requiring GDM, and pregestational diabetes). In other recent studies, it has been reported that fetal thymus sizes were smaller in the diet-controlled GDM, insulin-requiring GDM, and pregestational diabetes groups compared to the control group, with the lowest values observed in the pregestational diabetes group.[14] Contrary to these findings, our data revealed comparable TTR, AP diameter, and TTD across the insulin-requiring, diet-controlled, and healthy control cohorts.

This heterogeneity across the literature may be explained by two opposing pathophysiological forces acting concurrently on the thymic tissue. Fetal hyperinsulinemia and insulin-like growth factors, both triggered by maternal hyperglycemia, can stimulate somatic growth and cellular proliferation within the fetal thymus, an organ characterized by a high density of insulin receptors. Conversely, the chronic low-grade inflammation, oxidative stress, and heightened glucocorticoid exposure inherent to GDM can accelerate thymocyte apoptosis, driving structural involution. The equivalent thymic dimensions observed across our study cohorts may reflect a functional equilibrium between these counterbalancing anabolic and apoptotic pathways. Moreover, because our cohort comprised tightly controlled GDM patients and excluded pregestational diabetes with end-organ vascular disease, the development of a hostile intrauterine milieu severe enough to precipitate overt thymic regression was likely prevented.

Another critical consideration in fetal thymic assessment is the gestational timing of the sonographic evaluation. In our study, the decision to perform measurements between 28 and 36 weeks of gestation was guided by two primary rationales. First, the structural sequelae of GDM, namely fetal hyperinsulinemia, organomegaly, and subclinical inflammatory stress in the microenvironment, reach their most pronounced phenotypic expression during the third trimester. Second, technical constraints dictate this window: prior to 28 weeks, the acoustic contrast delineating thymic tissue margins remains ill-defined, whereas beyond 36 weeks, advancing ossification of the sternum and rib cage generates substantial acoustic shadowing that compromises anterior mediastinal visualization. In our cohort, thymic biometry was measured at a median of 32 gestational weeks across all study groups, with no significant intergroup differences in the timing of the examinations. This methodological uniformity minimizes the confounding effect of gestational age on our outcomes. Because absolute thymic dimensions expand with advancing gestation, comparing raw linear diameters across different gestational ages can introduce substantial bias. The principal clinical strength of the TTR, by indexing the thymic AP diameter to the intrathoracic diameter within the same plane, lies in its capacity to standardize individual variations in fetal biometry and somatic growth.

Notably, the thoracic AP diameter was larger in the insulin-requiring group, consistent with its higher abdominal circumference and birth weight, which reflect accelerated somatic growth. Although the thymic AP diameter also tended to be larger in this group (16.2 vs. 15.2 mm in controls), TTR remained comparable. This suggests that thymic growth in insulin-requiring GDM is proportional to thoracic growth rather than reflecting selective enlargement or involution, and it illustrates the advantage of using a ratio over absolute diameters when fetal size differs between groups.

The most striking finding from our TTR analysis lies not in biometric comparisons between cohorts, but in its association with neonatal outcomes. Although TTR values were comparable across the study groups, higher TTR measurements in the overall population correlated significantly with higher 1st- and 5th-minute Apgar scores. While the modest correlation coefficients suggest that this relationship is not strong, it nonetheless demonstrates a statistically consistent directionality. It is well established that the fetal thymus responds to intrauterine stress with rapid involution, and robust evidence from systematic reviews and meta-analyses has linked reduced thymic size to heightened neonatal morbidity, including preterm birth, chorioamnionitis, and early-onset neonatal sepsis.[15,16] Consequently, TTR appears to reflect the biological response of the fetus to the intrauterine environment

rather than serving as a disease-specific marker for any single maternal diagnosis.

Studies investigating the direct relationship between fetal thymic biometry and Apgar scores in pregnancies complicated by GDM remain scarce, and the available evidence is inconsistent. While several prospective cohort studies have linked GDM to an increased risk of low 1st and 5th minutes Apgar scores, others focusing on term neonates found no significant disparity in Apgar scores between diabetic and control cohorts.[17,18]

The absence of differences in Apgar scores between groups in our cohort is attributable to our study population consisting of pregnant women who achieved glycemic control through close antenatal monitoring and who were free of comorbidities. Therefore, the independent association observed between TTR and these scores, despite no difference in Apgar scores between groups, suggests that TTR may reflect individual fetal reserve and adaptive capacity rather than merely indicating GDM.

Although maternal BMI was demonstrably elevated in both GDM arms relative to controls in our sample, TTR remained similar across groups, suggesting that maternal obesity does not directly modify fetal TTR. A study by Yildirim et al.[19] reported an increase in the absolute dimensions of the fetal thymus in obese pregnant women, attributing this increase to metabolic and inflammatory processes; however, that study evaluated absolute thymus size rather than TTR. Since the relationship between BMI and TTR was not assessed by direct correlation analysis or multivariate models in our study, one should not conclude definitively that TTR is independent of BMI.

The strengths of our study include the prospective cohort design, the standardized diagnostic protocol, the explicit separation of GDM phenotypes by treatment modality (diet versus insulin), an adequate sample size confirmed by G*Power analysis, and the performance of all ultrasonographic measurements by the same experienced perinatologist using the same ultrasound system and standard anatomical planes, eliminating inter-observer variability.

Study Limitations

Several limitations merit acknowledgment. Although the enrolled sample met the pre-specified power requirement, the data clustered around higher scores rather than spanning a wide range due to the small number of cases involving severe neonatal asphyxia or markedly low Apgar scores (<7). Second, the thymus was assessed at a single time point during pregnancy, and serial measurements were not performed; consequently, changes in thymic size over time following the diagnosis and treatment of GDM could not be determined. Furthermore, the lack of data on HbA1c levels, fasting and postprandial blood glucose monitoring, and daily insulin requirements limits evaluation of the relationship between thymic dimensions and glycemic control. Finally, the assessment of neonatal well-being did not include, alongside Apgar scores, monitoring of umbilical artery blood pH, base deficit, the need for neonatal intensive care, or the incidence of neonatal hypoglycemia. Incorporating these factors into the analysis would provide a much clearer picture of the role of fetal TTR in predicting neonatal morbidity.

CONCLUSION

However, the finding that TTR was positively, albeit weakly, correlated with 1st- and 5th-minute Apgar scores suggests that fetal thymic size normalized to thoracic dimensions may be more informative regarding early neonatal adaptation than isolated diameters. TTR may, therefore, represent a candidate sonographic marker of the fetus's capacity to adapt to intrauterine stress, rather than a parameter guiding the diagnostic or therapeutic approach to GDM. Nevertheless, for TTR to be utilized as a clinical predictive tool, multicenter prospective studies involving serial screenings across diverse populations and incorporating objective cord blood parameters are required.

Ethics

Ethics Committee Approval: The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and approved by the Ethics Committee of University of Health Sciences Türkiye, İzmir Tepecik Training and Research Hospital (approval number: 2026/05-19; date: 04.06.2026).

Informed Consent: Prior to participation in the study, all participants were provided with detailed information, and written informed consent was obtained from each participant.

Footnotes

Authorship Contributions

Surgical and Medical Practices: R.T., Z.E.Ç., Concept: R.T., E.D., Design: R.T., E.D., Data Collection or Processing: R.T., S.T.C., Z.E.Ç., Analysis or Interpretation: R.T., E.D., S.T.C., Z.E.Ç., Literature Search: R.T., S.T.C., Z.E.Ç., Writing: R.T.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that this study received no financial support.

REFERENCES

1. Sweeting A, Wong J, Murphy HR, Ross GP. A clinical update on gestational diabetes mellitus. Endocr Rev. 2022;43:763-93.
2. Ornoy A, Becker M, Weinstein-Fudim L, Ergaz Z. Diabetes during pregnancy: a maternal disease complicating the course of pregnancy with long-term deleterious effects on the offspring. A clinical review. Int J Mol Sci. 2021;22:2965.
3. Benhalima K, Van Crombrugge P, Moyson C, et al. Characteristics and pregnancy outcomes across gestational diabetes mellitus subtypes based on insulin resistance. Diabetologia. 2019;62:2118-28.

4. Ansari AR, Liu H. Acute thymic involution and mechanisms for recovery. Arch Immunol Ther Exp (Warsz). 2017;65:401-20.
5. Ekin A, Gezer C, Taner CE, Solmaz U, Gezer NS, Ozeren M. Prognostic value of fetal thymus size in intrauterine growth restriction. J Ultrasound Med. 2016;35:511-7.
6. Kuyucu M, Atakul BK, Adıyaman D, Bayraktar B, Gölbaşı H, Pala H. Two- and three-dimensional ultrasonographic evaluation of the fetal thymus in women with preterm premature rupture of membranes: a prospective study. J Clin Obstet Gynecol. 2023;33:103-10.
7. Ye W, Luo C, Huang J, Li C, Liu Z, Liu F. Gestational diabetes mellitus and adverse pregnancy outcomes: systematic review and meta-analysis. BMJ. 2022;377:e067946.
8. International Association of Diabetes and Pregnancy Study Groups Consensus Panel; Metzger BE, Gabbe SG, Persson B, et al. International association of diabetes and pregnancy study groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy. Diabetes Care. 2010;33:676-82.
9. Kasap E, Saglam C, Torun R, et al. Fetal thymus size in pregnant women with Hashimoto disease. J Clin Ultrasound. 2024;52:1066-70.
10. Schober P, Boer C, Schwarte LA. Correlation coefficients: appropriate use and interpretation. Anesth Analg. 2018;126:1763-8.
11. Balkas G, Çelen Ş. Maternal systemic inflammation and fetal thymic size in diabetic pregnancies: predictive role of hematological biomarkers. J Clin Med. 2025;14:7619.
12. Mansouri N, Khezripour E, Rashtiani N, Bagherinia M, Azizi A. Gestational diabetes and its effect on fetal thymus size: a case-control study. BMC Pregnancy Childbirth. 2025;25:330.
13. Dörnemann R, Koch R, Möllmann U, et al. Fetal thymus size in pregnant women with diabetic diseases. J Perinat Med. 2017;45:595-601.
14. Sinaci S, Sahin D. Can diabetes influence fetal thymus size during pregnancy? J Obstet Gynaecol Res. 2023;49:1491-7.
15. Caissutti C, Familiari A, Khalil A, et al. Small fetal thymus and adverse obstetrical outcome: a systematic review and a meta-analysis. Acta Obstet Gynecol Scand. 2018;97:111-21.
16. Villamor-Martinez E, Lubach GA, Rahim OM, et al. Association of histological and clinical chorioamnionitis with neonatal sepsis among preterm infants: a systematic review, meta-analysis, and meta-regression. Front Immunol. 2020;11:972.
17. Yeagle KP, O'Brien JM, Curtin WM, Ural SH. Are gestational and type II diabetes mellitus associated with the Apgar scores of full-term neonates? Int J Womens Health. 2018;10:603-7.
18. Szymoniak K, Olejnik N, Zimny M, Sipak-Szmigiel O. Influence of gestational diabetes mellitus on the birth condition of full-term newborns - a preliminary study. Pomeranian J Life Sci. 2022;68:47-51.
19. Yildirim M, Ipek A, Dauletkazin G, et al. Sonographic measurement of the fetal thymus: relationship with maternal obesity. J Clin Ultrasound. 2017;45:277-81.