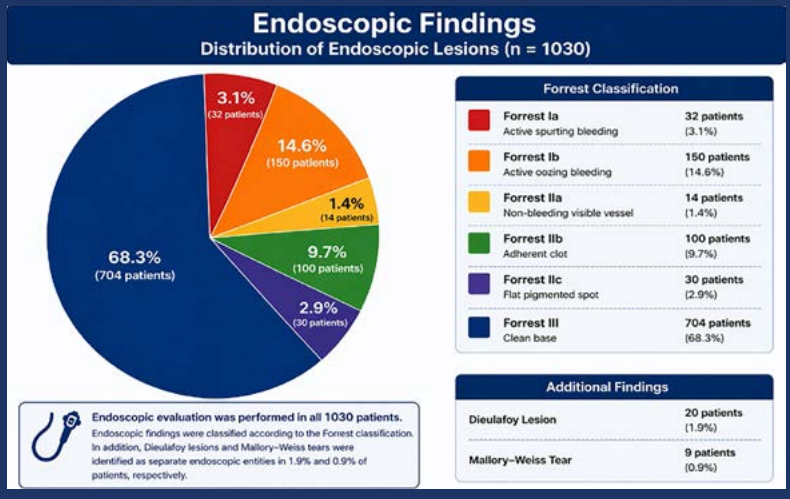
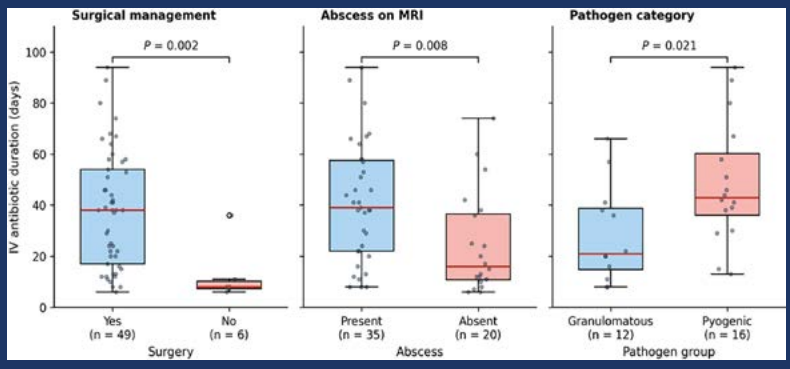
E-ISSN 2822-6771

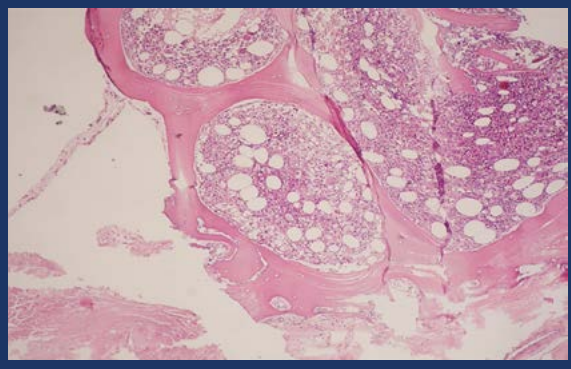
COMPREHENSIVE MEDICINE

Formerly İstanbul Kanuni Sultan Süleyman Tıp Dergisi

Volume 18 Supplement 1 October Year 2026







Rockall vs GBS in Non-Variceal Bleeding

Neurosurgical Spondylodiscitis Cohort from Türkiye

Clear Cell Carcinoma with Bone Marrow Metastasis

www.comprehensivemedicine.org

COMPREHENSIVE MEDICINE

Formerly İstanbul Kanuni Sultan Süleyman Tıp Dergisi

Volume 18 Supplement 1 October Year 2026

EDITORIAL BOARD

On behalf of University of Health Sciences,
İstanbul Kanuni Sultan Süleyman Training and Research
Hospital Head Physician

Prof. Serdar Çevik, M.D.

Editor-in-Chief

Erkan Somuncu
Department of General Surgery, Istanbul Kanuni Sultan Süleyman Training and Research Hospital, Istanbul, Turkiye
dr.somuncu@gmail.com

Associate Editors

Serkan Doğan
Department of Emergency Medicine, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
sercem2003@hotmail.com

Murat Akarsu
Department of Internal Medicine, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
muratakarsu79@gmail.com

Section Editors

Perinatology
Alev Aydın
Department of Gynecology and Obstetrics, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
alevatis8@gmail.com

Gynecologic Oncology
Merve Aldıkaçtıoğlu Talmaç
Department of Gynecologic Oncology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
drmrve@hotmail.com

Pediatrics
Himmet Haluk Akar
Department of Pediatrics, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
himmetakar@gmail.com

Pediatric Intensive Care
Nihal Akçay
Department of Pediatric Intensive Care, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
drnihalakcay@gmail.com

Cardiology
Bilal Çuğlan
Department of Cardiology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
bilalcuglan@hotmail.com

Internal Medicine
Şengül Aydın Yoldemir
Department of Internal Medicine, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
sengulaydinn@gmail.com

Neurology
Mesrure Köseoğlu
Department of Neurology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
mesrurekoseoglu@gmail.com

Neurosurgery
Serdar Çevik
Department of Neurosurgery, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
dr.serdarcevik@gmail.com

COMPREHENSIVE MEDICINE

Formerly İstanbul Kanuni Sultan Süleyman Tıp Dergisi

EDITORIAL BOARD

Volume 18 Supplement 1 October Year 2026

Section Editors

General Surgery

Yasin Kara

Department of General Surgery, Istanbul Kanuni Sultan Süleyman Training and Research Hospital, Istanbul, Turkiye
yasinkara32@windowslive.com

Orthopedics and Traumatology

Bülent Kılıç

Department of Orthopedics and Traumatology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
drbulentk@hotmail.com

Adem Şahin

Department of Orthopedics and Traumatology, Istanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
sahinadem17@gmail.com

Emergency Medicine

Ramiz Yazıcı

Department of Emergency Medicine, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
dr.ramiz.yazici@gmail.com

Anesthesiology and Resuscitation

Kadir Arslan

Department of Anesthesiology and Resuscitation, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
kadir.arslan@sbu.edu.tr

Eye Diseases

Sadık Etka Bayramoğlu

Department of Eye Diseases, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
sadiketka@windows.live.com

Nihat Sayın

Department of Eye Diseases, Istanbul Kanuni Sultan Süleyman Training and Research Hospital, Istanbul, Turkiye
nihatsayin@yahoo.com

Medical Biochemistry

Cemal Kazezoğlu

Department of Medical Biochemistry, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
cemal.kazezoglu@sbu.edu.tr

Medical Pathology

Taşkın Erkin Üresin

Department of Medical Pathology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
erkinuresin@hotmail.com

Radyology

Behice Kaniye Yılmaz

Department of Radiology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
behiceyilmaz@gmail.com

Infectious Diseases

Ayşe İnci

Department of Infectious Diseases and Clinical Microbiology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
ays.2004@yahoo.com.tr

Pediatric Neurology

Adnan Ayvaz

Department of Pediatric Neurology, Faculty of Medicine, İstanbul Yeniyuzyıl University, İstanbul, Türkiye
draayvaz@gmail.com

Pediatric Rheumatology

Fatma Gül Demirkan

Department of Pediatric Rheumatology, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
fatmagy@gmail.com

Statistical Editor

Türker Demirtakan

Department of Emergency Medicine, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
turkerkaya344@gmail.com

Mehmet Demir

Department of Neurology, Istanbul Kanuni Sultan Süleyman Training and Research Hospital, Istanbul, Turkiye
md.mehmetdemir@hotmail.com

COMPREHENSIVE MEDICINE

Formerly İstanbul Kanuni Sultan Süleyman Tıp Dergisi

Volume 18 Supplement 1 October Year 2026

AIMS AND SCOPE

Comprehensive Medicine is the official publication of Kanuni Sultan Süleyman Training and Research Hospital. It is published regularly, four times a year, in January, April, July and October as an electronic journal.

The journal was published by Bakırköy Obstetrics and Gynecology Hospital in 1990s as JOPP (gynecology obstetrics, pediatric and pediatric surgery) and has reached the status of 'peer-reviewed journal' in 2009. The journal was changed to İstanbul Kanuni Sultan Süleyman Tıp Dergisi in 2014 and known as Comprehensive Medicine beginning from 2023.

The language of the journal is English. It is a peer-reviewed journal and does not charge any fee from readers or authors. All published articles are open access.

The aim of Comprehensive Medicine is to provide students, who are educated in the field of medicine in all fields of medicine, and researchers to all physicians who are in the field of health and information about disease treatment and prevention methods. The journal publishes original research articles including gynecology, obstetric, pediatrics, pediatric surgery and basic medical sciences, case reports, reviews and announcements of panels and congresses. Comprehensive Medicine is to publish original researches on international level qualified, continuous and general medicine and periodically evaluated at the highest level both clinically and scientifically according to the konusunda double-blind, system for the consultant and writer. In addition to this, they include collections covering basic innovations about education, letters to the editor, and case reports.

Moreover, by conducting an objective evaluation within the framework of the ethical rules of the journal, the consultant aims to enable the reviewers to transfer their experiences to the authors, to give the author a perspective on how to write a scientific article, and to contribute to the elimination of the scientific value of the articles published in our country.

Publication History

Currently: Comprehensive Medicine | E-ISSN: 2822-6771 (2023 - current)

Formerly: Medical Journal of Istanbul Kanuni Sultan Suleyman / İstanbul Kanuni Sultan Süleyman Tıp Dergisi | ISSN: 2148-273X / e-ISSN: 2667-7458 (2014-2022).

Formerly: JOPP (Gynecology Obstetrics Pediatrics and Pediatric Surgery Journal) | ISSN: 1308-6715 (until 2013)

Comprehensive Medicine is indexed in TÜBİTAK ULAKBİM (TR Dizin), EBSCO, DOAJ, Türk Atıf Dizini (Turkish Citation Index), J-Gate, Türk Medline, GALE, CAS, CEEAS, OUCI, Scilit, WorldCat and ASCI and Embase.

Open Access Policy

This journal provides immediate open access to its content on the principle that making research freely available to the public supports a greater global exchange of knowledge.

Open Access Policy is based on the rules of the Budapest Open Access Initiative (BOAI) http://www.budapestopenaccessinitiative.org/. By "open access" to peer-reviewed research literature, we mean its free availability on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of these articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The only constraint on reproduction and distribution, and the only role for copyright in this domain, should be to give authors control over the integrity of their work and the right to be properly acknowledged and cited.

As publisher of the Comprehensive Medicine, Kare Media House does not demand any subscription fee, publication fee or similar payment for the access into electronic resources.

Permission Requests

For all published articles, Comprehensive Medicine accepts the "Creative Commons Attribution License" (Attribution 4.0 International CC BY)

This license lets others distribute, remix, adapt, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

Open access is an approach that reinforces interdisciplinary development and encourages cooperation between different disciplines. For this reason, Comprehensive Medicine Journal contributes to its field by offering more access to its articles and a more transparent review process.

Advisory Board: https://comprehensivemedicine.org/advisory_board

Writing Guide: https://comprehensivemedicine.org/instructions-to-authors

Ethic and Policies: https://comprehensivemedicine.org/ethics-policies

OPEN ACCESS

Advertisement / Publisher Corresponding Address

For requests concerning advertising, please contact the Publisher:

Kare Medya İletişim Hizmetleri Tic. Ltd. Şti.

Publications Coordinator: Ahmet Keskin

Graphic Design: Duygu Şimşek

Address: Göztepe Mah. Fahrettin Kerim Gökay Cad. No: 200 D: 2 Göztepe, Kadıköy, İstanbul-Türkiye

Telephone: +90 (216) 550 61 11 - **Fax:** +90 (216) 550 61 12

Web page: www.karepb.com - **E-mail:** kare@karepb.com

Online Publication Date: October 2026 ISSN: 2148-273X E-ISSN: 2667-7458

International periodical journal published four times in a year.

COMPREHENSIVE MEDICINE

Formerly İstanbul Kanuni Sultan Süleyman Tıp Dergisi

CONTENTS

Volume 18 Supplement 1 October Year 2026

RESEARCH ARTICLES

1 Comparative Correlation of Rockall and Glasgow-Blatchford Scoring Systems with One-Month Mortality and Rebleeding in Non-Variceal Upper Gastrointestinal Bleeding
Nezihe Berrin Dodur Önalan, Mahmut Ozan Aydın, Cenk Özkan, Halil Alış

11 Predictive Value of Pretreatment Inflammation-Based Prognostic Scores for Primary Non-Muscle-Invasive Bladder Cancer
Haydar Güler, Uğur Yücetaş, Erkan Sönmezay, Erkan Erkan, Yaşar Pazır, Mustafa Kadıhasanoğlu

16 The Effects of Mild and Moderate Hypothermia on Liver Regeneration, Liver Damage, and Apoptosis in Hepatectomized Rats
Ali Önder Devay, Mehmet Özdoğan, Pamir Eren Ersoy, Rıza Haldun Gündoğdu, Recep Aktimur, Selma Göğkuş Beyeç, Seda Duygulu Devay, Neşe Lortlar Ünlü

25 Phenotypic Characteristics and Co-sensitization Patterns in Pediatric *Cupressus Arizonica* Sensitization
Hande Yüksel Bulut, Enes Çelik

32 Evaluation of Physicians' Knowledge and Proficiency in The Management of Anaphylaxis and The Use of Adrenaline Auto-Injectors
Mehmet Özkaya, Hatice Çelik Tuğlu

38 Intraoperative Blood Loss Is a Surrogate of Surgical Complexity Rather Than an Independent Predictor of Pedicle Screw Loosening: A Matched Cohort Analysis
Evren Sönmez, Lokman Ayhan, Harun Duran

47 Frequency of Sarcopenia and Associated Factors in Older Adults with Non-dialysis Dependent Chronic Kidney Disease: A Cross-sectional Study
Tuğçe Emiroğlu Gedik, Onur Kağan Öztaş, Mustafa Ortaboz, Aylin Şendoğan, Murat Akarsu

53 Many Videos Demonstrate the Operation, Few Teach the Operation: Educational Completeness and Workflow Representation in Oncologic Laparoscopic Adrenalectomy Videos on YouTube
Adem Özcan, Gizem Güneş, Mahmut Ozan Aydın, İhsan Barış Müldür

67 Long-Term Anal Fistula Formation After Incision and Drainage of Cryptoglandular Perianal Abscess: A Single-Center Retrospective Cohort Study with 10-Year Follow-up
Emre Özoran, Tutku Tüfekçi

74 Posttraumatic Growth and Depreciation After Childbirth: The Role of Mode of Delivery and Psychological Resilience in a Turkish Postpartum Sample
Berivan Güzelbağ, Melike Punduk Yılmaz, Filiz Yarşılıkal Güleroğlu, Ali Çetin

84 Integration of RES-Q-Compatible Stroke Documentation into the Hospital Information System: Evaluation of Stroke Center Indicators
Hatice Yelda Yıldız, Umut Esen, Yavuz Bekmezci, Neslihan Oral, Nihal Ünlütürk, Ferda Kaya Zaman, Tülay Özer, Şerif Köksal, Yakup Krespi

94 Association of Lithium Treatment with Biological Rhythms in Euthymic Patients with Bipolar Disorder
Ayşe Kurtulmuş, Fatma Büşra Parlakkaya, Halenur Şamandar, Pınar Ertunç, Zülal Çelik

103 Clinical, Microbiological and Treatment Characteristics of Spondylodiscitis: A Six-Year Single-Centre Neurosurgery Experience from a Brucellosis-Endemic Region of Türkiye
Salim Katar, Adil Can Karaoğlu, Muhammed Fatih Sarı, Mehmet Ali Tüz, Rukiye İnan Sarıkaya, Emre Fidan, Hüseyin Utku Adilay

112 Admission Neutrophil-to-Lymphocyte Ratio and Hospital Stay After Surgery for Chronic Subdural Hematoma Outcomes
Fatih Çalış, Mohammed Aladdam, Mehmet Ali Kahraman, Said Abumarasa, Ahmet Faruk Özdemir, Semih Naci Balak

119 Patient-Reported Outcomes Following Transsphenoidal Surgery for Acromegaly: A Cross-Sectional Study of Quality of Life, Sleep Quality, and Body Image
Ozan Başkurt, Nurperi Gazioğlu

CASE REPORT

126 Atypical Presentation: A Clear Cell Carcinoma Case Diagnosed with Bone Marrow Metastasis
Dilek Keskin, Ahmet Can Toksöz, Taşkın Erkinüresin, Murat Akarsu

DOI: 10.14744/cm.2026.20082
Compreh Med 2026;18(Suppl. 1):1-10

RESEARCH ARTICLE

Comparative Correlation of Rockall and Glasgow-Blatchford Scoring Systems with One-Month Mortality and Rebleeding in Non-Variceal Upper Gastrointestinal Bleeding

Nezihe Berrin Dodur Önalan[1], Mahmut Ozan Aydın[1], Cenk Özkan[1], Halil Alış[2]

[1]Department of General Surgery, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
[2]Department of General Surgery, İstanbul Aydın University Faculty of Medicine, İstanbul, Türkiye

ABSTRACT

Objective: Upper gastrointestinal bleeding (UGIB) is a common and potentially life-threatening emergency. Risk stratification using composite scoring systems is essential for guiding clinical management. This study evaluated the correlations of the Rockall and Glasgow-Blatchford scoring (GBS) systems with one-month mortality and rebleeding in patients with non-variceal UGIB and compared their performance with individual clinical and laboratory parameters.

Materials and Methods: A retrospective analysis was performed using prospectively collected data from 1,030 patients with confirmed non-variceal UGIB admitted to a tertiary referral center between January 2001 and September 2005. Rockall and GBS scores were calculated for each patient. Correlations between clinical, laboratory, and endoscopic variables and study outcomes were assessed using Spearman's correlation analysis.

Results: The mean age was 54.1 years, and 69.8% of the patients were male. Follow-up data were available for 262 of the 1,030 patients because a considerable proportion of the cohort could not be reached; therefore, rebleeding and mortality analyses were conducted within this subset. Both scoring systems demonstrated statistically significant correlations with one-month mortality, with the Rockall score showing a stronger association (r=0.507 vs. r=0.425). Both scores also correlated with rebleeding (r=0.259 and r=0.247, respectively). Among individual parameters, hemodynamic stability showed the strongest correlation with mortality (r=−0.346), followed by age (r=0.332). Both systems showed stronger correlations than individual variables.

Conclusion: The Rockall score demonstrates a stronger correlation with mortality than the GBS in patients with non-variceal UGIB. Both systems showed stronger correlations than individual clinical parameters, supporting structured multivariable risk assessment. Hemodynamic stability and age are the most informative individual predictors of mortality.

Keywords: Glasgow-Blatchford score, non-variceal bleeding, one-month mortality, rebleeding, rockall score, upper gastrointestinal bleeding

How to cite this article: Dodur Önalan NB, Aydın MO, Özkan C, Alış H. Comparative Correlation of Rockall and Glasgow-Blatchford Scoring Systems with One-Month Mortality and Rebleeding in Non-Variceal Upper Gastrointestinal Bleeding. Compreh Med 2026;18(Suppl. 1):1-10.

INTRODUCTION

Upper gastrointestinal bleeding (UGIB) represents one of the most common gastrointestinal emergencies encountered in clinical practice and is associated with significant morbidity and mortality despite continued advances in endoscopic and pharmacological management. Early and accurate risk stratification at the time of admission is paramount, as it directly influences clinical decision-making, the timing of endoscopic intervention, and the appropriate allocation of hospital resources. To this end, several composite scoring systems have been developed and validated over the past decades. Among these, the Rockall and Glasgow-Blatchford scores (GBS) have gained the widest clinical acceptance, with the former uniquely incorporating endoscopic findings alongside clinical and demographic parameters, whereas the latter relies solely on variables obtainable at the bedside before endoscopy.

Address for Correspondence: Mahmut Ozan Aydın, Department of General Surgery, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
E-mail: op.dr.mahmutozanaydin@gmail.com **ORCID ID:** 0000-0002-7061-6957

Received date: 30.04.2026
Revised date: 06.05.2026
Accepted date: 02.06.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Despite their widespread use, the relative merits of these two scoring systems—and their performance compared with individual clinical and laboratory parameters—remain subjects of ongoing investigation. In this study, we aimed to evaluate the clinical correlations of the Rockall and GBS systems alongside a broad range of individual variables in patients with confirmed non-variceal UGIB, with the primary objective of identifying the most reliable predictors of one-month mortality and rebleeding.

MATERIALS and METHODS

This study was conducted between January 2001 and September 2005 in the Department of General Surgery at Bakırköy Dr. Sadi Konuk Training and Research Hospital. It represents a retrospective analysis of prospectively collected data. Ethics committee approval was obtained from Istanbul University of Health Sciences Kanuni Sultan Suleyman Training and Research Hospital Scientific Research Ethics Committee (Approval Number: 3, Date: 08.01.2025) and was conducted in accordance with the principles of the Declaration of Helsinki. Due to the retrospective nature of the study, the requirement for written informed consent was waived by the institutional ethics committee. Patients presenting with hematemesis, melena, or hematochezia were initially evaluated for suspected UGIB. The diagnosis of UGIB was established based on clinical findings, including rectal examination and nasogastric aspirate analysis, as well as endoscopic findings. Patients without available endoscopic data and those with endoscopically confirmed variceal bleeding were excluded. Following these exclusions, a total of 1,030 patients with confirmed non-variceal UGIB were included in the study.

Endoscopic evaluation was performed in all patients included in the study. Emergency endoscopy was performed in patients with ongoing bleeding and hemodynamic instability, whereas hemodynamically stable patients underwent endoscopy as early as feasible, generally within 24–48 hours. Endoscopic procedures were performed using a Pentax ESK-1000 videoendoscope, and findings were classified according to the Forrest classification. Patients with Forrest Ia, Ib, and IIa lesions underwent endoscopic injection therapy (sclerotherapy with 1:100,000 adrenaline injected into four quadrants surrounding the bleeding site). A second-look endoscopy was performed after 24 hours, and repeat injection therapy was administered in cases of rebleeding. Patients who failed endoscopic treatment and required ≥5 units of blood transfusion within 24 hours or remained hemodynamically unstable underwent emergency surgery. Patients were scheduled for repeat endoscopy at four weeks, when feasible, to assess ulcer healing and Helicobacter pylori status, during which morbidity, rebleeding, and mortality outcomes were evaluated. Eradication therapy was initiated when indicated. During hospitalization, patients received proton pump inhibitors or H2-receptor antagonists and were discharged with recommendations to continue medical therapy and discontinue nonsteroidal anti-inflammatory drugs and aspirin when applicable.

At admission, all patients were evaluated using two validated risk scoring systems, the Rockall score and GBS, to assess the risk of rebleeding and mortality. The Rockall score was calculated based on clinical and endoscopic parameters, including patient age, hemodynamic status (presence of hypotension and tachycardia), comorbid diseases, endoscopic diagnosis, and stigmata of recent hemorrhage observed during endoscopy. Detailed components and score allocations of the Rockall system are presented in Table 1.[1]

Table 1. Rockall scoring system for risk stratification in patients with UGIB

Rockall scores	Score
Age (years)	
<60	0
60–79	1
≥80	2
Hemodynamic status	
>100	1
Heart rate (bpm)	
Systolic blood pressure (mmHg)	
<100	2
Comorbidities	
Heart failure, ischemic heart disease	2
Renal failure, hepatic failure, metastatic cancer	3
Endoscopic signs (diagnostic)	
No lesion or Mallory–Weiss tear	0
Peptic ulcer, erosive disease, esophagitis	1
Cancer of the upper gastrointestinal tract	2
Endoscopic signs (hemorrhagic)	
Clean-base ulcer or flat, pigmented spot	0
Visible blood, active bleeding, visible vessel, adherent clot	2

The total score ranges from 0 to 11. Patients are classified as low risk (≤2), moderate risk (3–4), or high risk (≥5).[1] UGIB: Upper gastrointestinal bleeding

Table 2. Glasgow-Blatchford scoring system for the assessment of rebleeding and mortality risk in patients with UGIB

Glasgow-Blatchford Scores	Threshold	Score
Blood urea nitrogen (mmol/L)	6.5–7.9	2
	8.0–9.9	3
	10.0–24.9	4
	≥25.0	6
Hemoglobin for men (g/dL)	12.0–13.0	1
	10.0–11.9	3
	<10.0	6
Hemoglobin for women (g/dL)	10.0–12.0	1
	<10.0	6
Systolic blood pressure (mmHg)	100–109	1
	90–99	2
	<90	3
Heart rate (bpm)	>100	1
Melena	Present	1
Syncope	Present	2
Hepatic disease	Present	2
Cardiac failure	Present	2

The total score ranges from 0 to 23. Higher scores indicate increased risk and a greater need for clinical intervention.[2] UGIB: Upper gastrointestinal bleeding

Each component was assessed systematically, and the total Rockall score was recorded for each patient. According to the Rockall scoring system, patients were classified as low risk (0–2), moderate risk (3–4), or high risk (≥5).

GBS was calculated using clinical and laboratory variables obtained at presentation, including blood urea nitrogen, hemoglobin level, systolic blood pressure, and heart rate, as well as the presence of melena, syncope, hepatic disease, and cardiac failure. Urea values were converted to the corresponding blood urea nitrogen categories required for GBS calculation. Detailed components and score allocations of the Glasgow-Blatchford system are shown in Table 2.[2] All parameters were recorded at admission, and the total Glasgow-Blatchford score was determined for each patient.

The threshold for hypotension was defined as an arterial blood pressure below 90/60 mmHg, and tachycardia was defined as a heart rate greater than 100 beats per minute. Any death occurring within 30 days of admission—either during hospitalization or during the follow-up period after discharge—was recorded as a mortality event. Deaths were further categorized as UGIB-related or non-UGIB-related. Any recurrence of hematemesis, melena, or hematochezia during the one-month follow-up period was recorded as rebleeding.

Statistical Analysis

Statistical analyses were performed using SPSS software (IBM SPSS Statistics, version 25.0; IBM Corp., Armonk, NY, USA). Continuous variables were summarized as mean±SD (range), as appropriate, and categorical variables were presented as numbers and percentages. Correlations between clinical variables and study outcomes, including rebleeding and one-month mortality, were assessed using Spearman's correlation analysis. A p-value of <0.05 was considered statistically significant.

RESULTS

Study Group

A total of 1,030 patients were included in the study, of whom 719 (69.8%) were male and 311 (30.2%) were female. The mean age was 54.1 years (range: 12–90 years). Comorbidities were classified as cardiac, neoplastic, renal, cerebral, diabetes mellitus, respiratory, and hepatic diseases. The distribution of comorbidities was as follows: cardiac disease in 81 patients (7.9%), neoplastic disease in 24 (6.8%), renal disease in 22 (6.3%), cerebral disease in 21 (6.0%), diabetes mellitus in 20 (5.7%), respiratory disease in 19 (5.4%), and hepatic disease in 9 (2.6%).

Endoscopic evaluation was performed in all 1,030 patients. The mean time to endoscopy was 34.3 hours (range: 3–48 hours). Endoscopic findings were classified according to the Forrest classification. Accordingly, Forrest Ia lesions were observed in 32 patients (3.1%), Ib lesions in 150 (14.6%), IIa lesions in 14 (1.4%), IIb lesions in 100 (9.7%), IIc lesions in 30 (2.9%), and III lesions in 704 (68.3%). Lesions not included in the Forrest classification, such as Dieulafoy lesions and Mallory–Weiss tears, were recorded separately and were not included in the Forrest distribution. These lesions were identified in 1.9% and 0.9% of patients, respectively. The distribution of endoscopic findings is presented in Figure 1.

At admission, 891 patients (86.5%) were hemodynamically stable, whereas 58 (5.6%) were hypertensive, 54 (5.2%) were tachycardic, 16 (1.6%) were hypotensive, 8 (0.8%) were hypotensive and tachycardic, 2 (0.2%) were hypertensive and tachycardic, and 1 (0.1%) was bradycardic. The mean systolic blood pressure was 114.9 mmHg (range: 50–210), the mean diastolic blood pressure was 69.9 mmHg (range: 30–120), and the mean heart rate was 87 beats per minute (range: 55–140).

Laboratory parameters were as follows: hemoglobin, 9.3 g/dL (range: 3.01–16.3); hematocrit, 27.5% (range: 8.04–47.4);

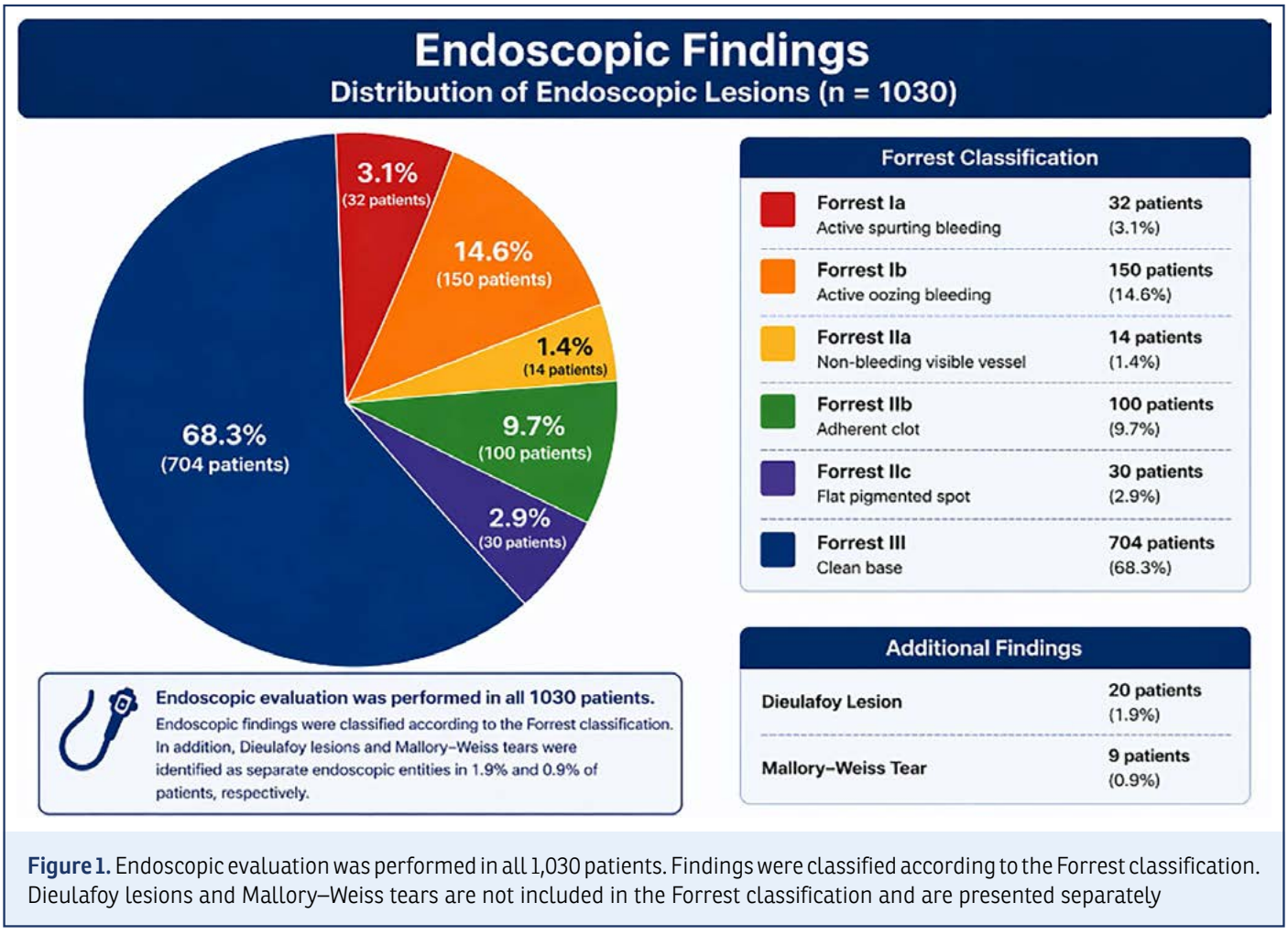


Figure 1. Endoscopic evaluation was performed in all 1,030 patients. Findings were classified according to the Forrest classification. Dieulafoy lesions and Mallory–Weiss tears are not included in the Forrest classification and are presented separately

and urea, 76.0 mg/dL (range: 8–633). The mean number of blood transfusions administered was 1.65 units (range: 1–12). Among the 1,030 patients, 21 (2.03%) required surgical intervention because of hemodynamic instability or failure to respond to endoscopic treatment.

Of the total cohort of 1,030 patients, follow-up data were available for 262. Among these, 232 patients (88.9%) did not experience rebleeding within the first month after discharge, whereas 30 patients (11.1%) developed recurrent bleeding during this period. During the one-month follow-up period, a total of 62 deaths were recorded. Of these, 36 deaths were directly attributable to UGIB, whereas 26 were due to non-UGIB-related causes. UGIB-related mortality occurred predominantly in the high-risk group according to the Rockall classification, with 35 of 36 deaths occurring in this group, whereas only one UGIB-related death occurred in the low-risk group and none occurred in the moderate-risk group. One-month mortality and rebleeding rates in the study cohort are presented in Figure 2.

The mean Rockall score was 2.45±1.99, whereas the mean Blatchford score was 7.15±4.49. Based on the Rockall scoring system, patients were stratified into low-, moderate-, and high-risk categories. The distribution of patients according to individual Rockall scores is presented in Figure 3, whereas the distribution across Rockall risk groups is shown in Figure 4. In addition, the distribution of patients according to Blatchford scores is shown in Figure 5, providing complementary information on clinical risk at presentation.

The relationships between UGIB-related mortality and various clinical, laboratory, and management-related variables were evaluated using Spearman's correlation analysis. A wide range of parameters, including demographic characteristics, hemodynamic findings at admission, laboratory values, endoscopic findings, treatment modalities, and risk scoring systems, were analyzed to determine their potential associations with mortality.

A positive correlation was observed between patient age and mortality ($r=0.332$, $p<0.05$). A significant association between

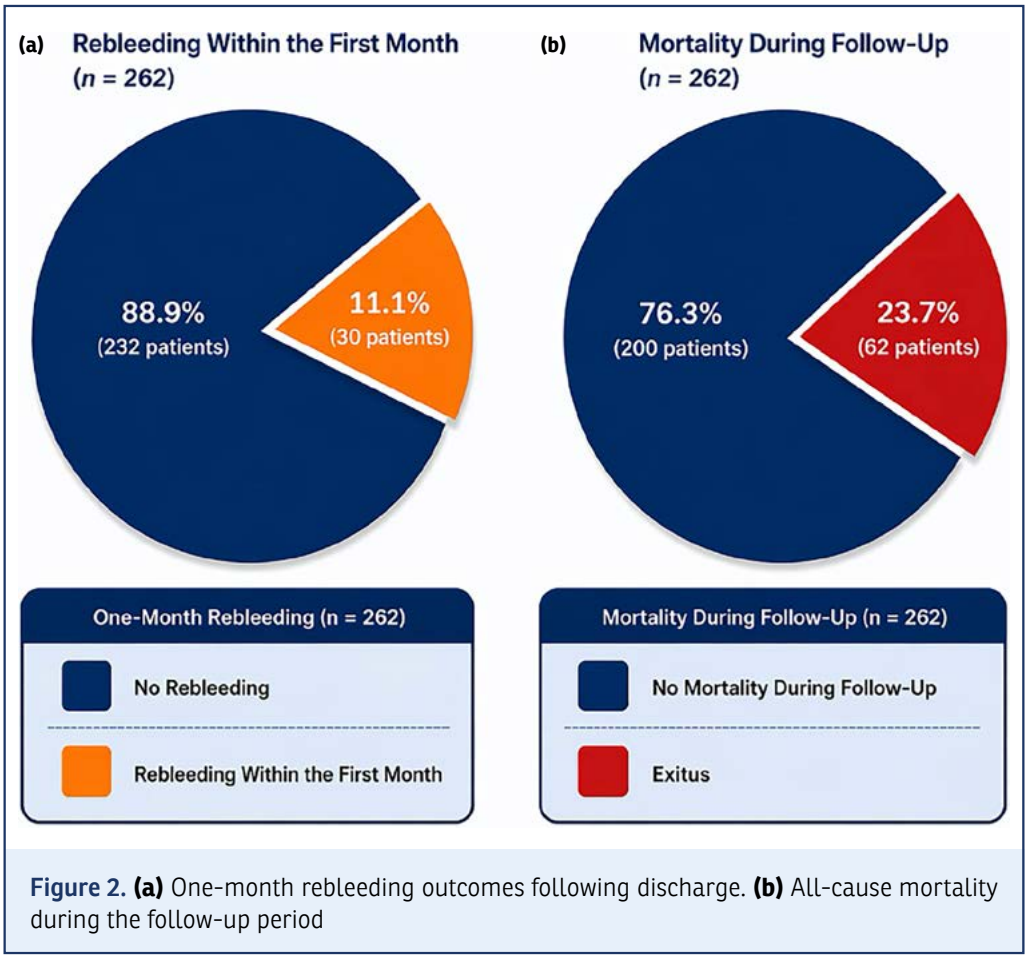


Figure 2. **(a)** One-month rebleeding outcomes following discharge. **(b)** All-cause mortality during the follow-up period

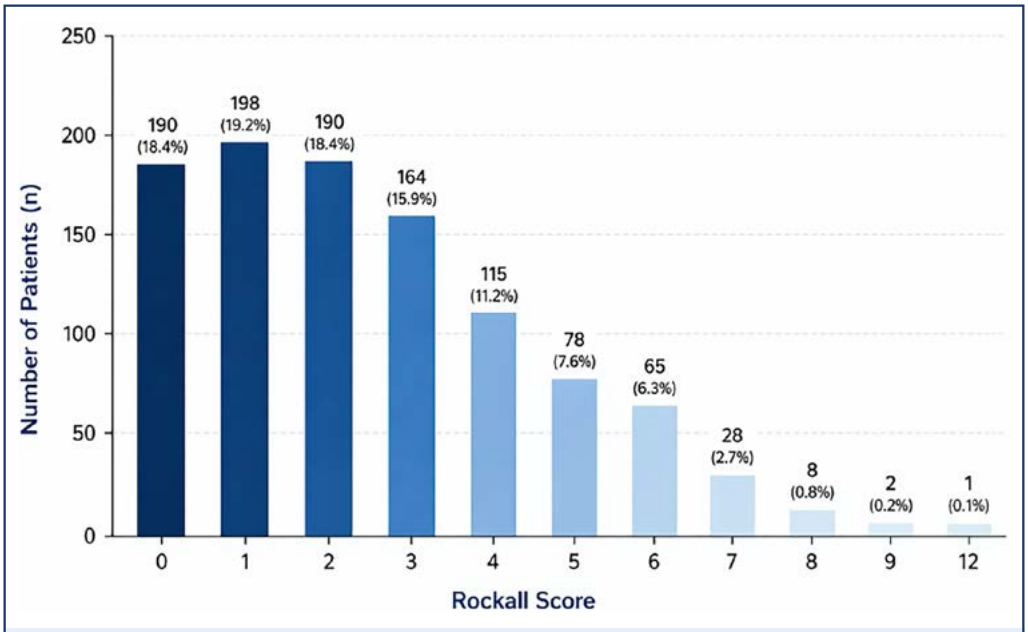


Figure 3. Distribution of patients according to Rockall score values in the study cohort (n=1,030)

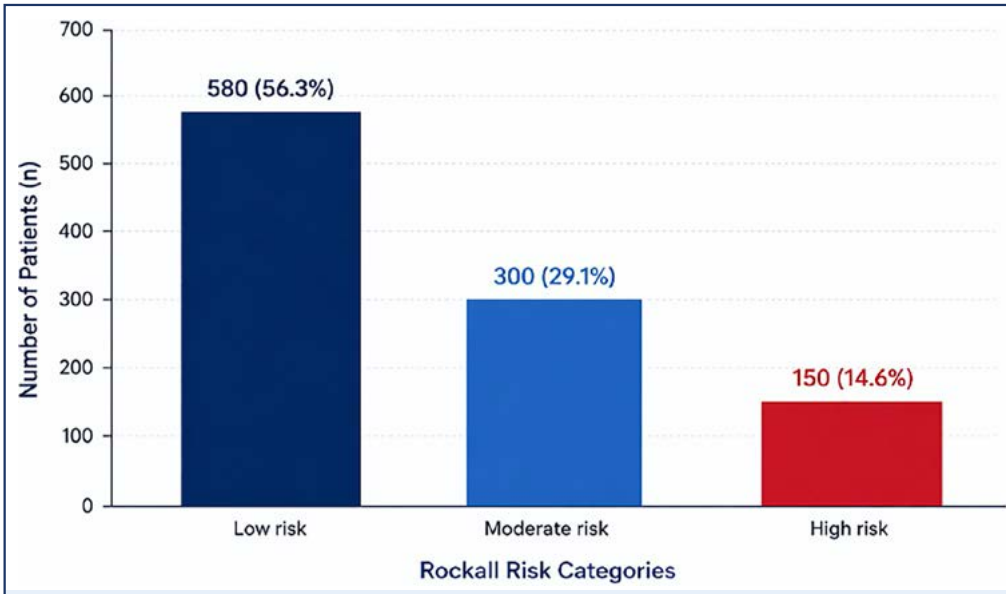


Figure 4. Distribution of patients according to Rockall risk categories (low, moderate, and high risk) in the study cohort (n=1,030)

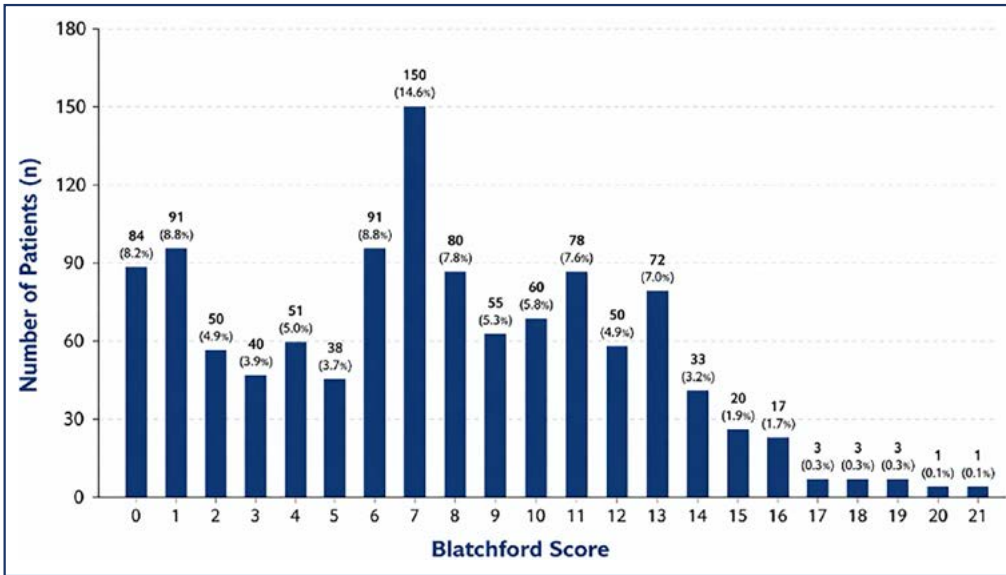


Figure 5. Distribution of patients according to Blatchford score values in the study cohort (n=1,030)

endoscopic findings and mortality was also identified; however, the strength of this correlation was weak (r=0.170, p<0.05).

Both systolic and diastolic blood pressure demonstrated weak inverse correlations with mortality (r=−0.147 and r=−0.152, respectively; p<0.05 for both). In contrast, hemodynamic stability showed a stronger inverse correlation (r=−0.346, p<0.05), indicating that hemodynamic stability was associated with a lower risk of mortality.

Hemoglobin and hematocrit levels were weakly and inversely associated with mortality (r=−0.179 and r=−0.159, respectively; p<0.05), whereas the urea level demonstrated a more pronounced positive correlation (r=0.286, p<0.05).

Both Rockall and Blatchford scores showed moderate correlations with mortality (p<0.05). Notably, these composite risk scoring systems—derived from several of the aforementioned clinical and laboratory parameters—demonstrated

Table 3. Spearman correlation analysis between clinical variables and mortality in patients with UGIB

Variable	Spearman's r	Direction	p
Age	0.332	Positive	<0.05
Endoscopic findings	0.170	Positive	<0.05
Systolic blood pressure	-0.147	Negative	<0.05
Diastolic blood pressure	-0.152	Negative	<0.05
Hemodynamic stability	-0.346	Negative	<0.05
Hemoglobin	-0.179	Negative	<0.05
Hematocrit	-0.159	Negative	<0.05
Urea	0.286	Positive	<0.05
Rockall score	0.507	Positive	<0.05
Blatchford score	0.425	Positive	<0.05
Sex	—	—	>0.05
Timing of endoscopy	—	—	>0.05
History of medication use	—	—	>0.05
Heart rate	—	—	>0.05
Blood transfusion	—	—	>0.05

Table 4. Spearman correlation analysis between clinical variables and rebleeding in patients with UGIB

Variable	Spearman's r	Direction	p
Age	—	—	>0.05
Endoscopic findings	0.288	Positive	<0.05
Systolic blood pressure	—	—	>0.05
Diastolic blood pressure	—	—	>0.05
Hemodynamic stability	—	—	>0.05
Hemoglobin	−0.126	Negative	<0.05
Hematocrit	—	—	>0.05
Urea	—	—	>0.05
Rockall score	0.259	Positive	<0.05
Blatchford score	0.247	Positive	<0.05
Sex	—	—	>0.05
Timing of endoscopy	—	—	>0.05
History of medication use	−0.197	Negative	<0.05
Heart rate	—	—	>0.05
Blood transfusion	—	—	>0.05

UGIB: Upper gastrointestinal bleeding

stronger associations than individual variables. Among all evaluated factors, the Rockall and Blatchford scores exhibited the strongest correlations with mortality, with the correlation being stronger for the Rockall score (r=0.507) than for the Blatchford score (r=0.425).

In contrast, no significant association was observed between mortality and sex, timing of endoscopy, history of medication use, heart rate, or blood transfusion (p>0.05). All correlation results are summarized in Table 3.

The relationships between rebleeding and various clinical variables were evaluated using Spearman's correlation analysis. The analysis demonstrated a significant association between endoscopic findings and the occurrence of rebleeding (r=0.288, p<0.05). A weak inverse correlation was observed between a history of medication use and rebleeding (r=−0.197, p<0.05), indicating a lower rate of rebleeding among patients receiving proton pump inhibitor therapy. Similarly, hemoglobin level demonstrated a weak inverse association with rebleeding (r=−0.126, p<0.05).

In contrast, no significant association was found between rebleeding and age, sex, timing of endoscopy, systolic or diastolic blood pressure, heart rate, hemodynamic stability, hematocrit level, urea level, or blood transfusion (p>0.05).

When evaluated using risk scoring systems, both Rockall and Blatchford scores showed weak but statistically significant correlations with rebleeding (r=0.259 and r=0.247, respectively; p<0.05). All correlation results are summarized in Table 4.

DISCUSSION

UGIB remains one of the most common and potentially life-threatening emergencies in surgical practice, carrying considerable morbidity and mortality despite advances in endoscopic and pharmacological management. Early and accurate risk stratification at the time of admission is therefore paramount, as it directly influences clinical decision-making, the timing of endoscopic intervention, and the appropriate allocation of hospital resources. Over the past decades, numerous scoring systems have been introduced to address this clinical need, among which the Rockall and GBS have gained the widest clinical acceptance and validation. In this study, we evaluated the clinical correlations of these two scoring systems alongside a broad range of individual clinical and laboratory variables in patients with confirmed non-variceal UGIB. Our findings demonstrate that both composite scoring systems showed statistically significant correlations with one-month mortality and rebleeding, with the Rockall score exhibiting a stronger association with mortality than the GBS.

In a multicenter study evaluating 2,977 patients across nine centers in China, Lu et al.[3] compared the clinical utility of the Rockall and GBS systems in patients with non-variceal UGIB. ROC analysis demonstrated that the Rockall score had superi-

or predictive value for mortality compared with the GBS (AUC: 0.804 vs. 0.622), whereas both scoring systems showed comparably limited performance in predicting rebleeding (AUC: 0.548 vs. 0.528). These findings were further corroborated by Custovic et al.,[4] who reported similar results in a single-center study of 237 patients conducted in Bosnia and Herzegovina, again demonstrating the superiority of the Rockall score in predicting mortality (AUC: 0.806 vs. 0.750). Notably, this study did not assess rebleeding as a separate endpoint. Furthermore, because the cohort included patients with variceal bleeding, the results should be interpreted with caution given the known differences in clinical course and scoring-system performance between variceal and non-variceal etiologies.

In contrast, Stanley et al.[5] reported comparable mortality prediction performance for both scoring systems in a multicenter prospective study of 1,555 patients across four UK centers, with no statistically significant difference between the full Rockall score and GBS (AUC: 0.790 vs. 0.804, p=0.20). Of particular note, this study was coauthored by Blatchford himself, lending considerable methodological authority to its conclusions. The mixed etiology of the cohort, which included both variceal and non-variceal bleeding, may have attenuated the observed differences between the two scoring systems. Importantly, however, this study highlighted a clinically meaningful distinction between the two scores: the GBS demonstrated markedly superior performance in predicting the need for blood transfusion (AUC: 0.932 vs. 0.752), underscoring the complementary rather than interchangeable nature of these two instruments—the Rockall score appearing better suited for mortality risk stratification, whereas the GBS offers greater utility in predicting the need for clinical intervention and blood transfusion, making it a more practical tool for early triage and resource allocation. Consistent with these findings, a meta-analysis by Kozai et al.[6] incorporating seven studies and 755 patients with non-variceal UGIB similarly demonstrated no statistically significant difference between the two scoring systems in predicting mortality (WMD=0.01, 95% CI: −0.06 to 0.08) or rebleeding (WMD=0.04, 95% CI: −0.03 to 0.11). However, the GBS was superior in predicting the need for blood transfusion (WMD=0.09, 95% CI: 0.01–0.16) and surgical intervention (WMD=0.21, 95% CI: 0.14–0.29).

In the present study, the Rockall score demonstrated a stronger correlation with one-month mortality than the GBS (r=0.507 vs. r=0.425). These discrepancies may reflect the underlying structural differences between the two systems: as noted by Custovic et al.,[4] mortality in UGIB largely reflects the burden of underlying comorbidities rather than the bleeding episode itself—parameters explicitly incorporated into the Rockall score but absent from the Glasgow-Blatchford score. The latter, which incorporates hemoglobin and hemodynamic variables, appears better suited to predicting the immediate clinical demands of the bleeding episode, including the need for transfusion and endoscopic or surgical intervention.

In a landmark international consensus statement developed in 2010 by a multidisciplinary panel of 34 experts from 15 countries—the International Consensus Upper Gastrointestinal Bleeding Conference Group—Barkun et al.[7] emphasized that composite prognostic scoring systems are recommended over individual clinical parameters for the early stratification of patients with non-variceal UGIB into low- and high-risk categories for rebleeding and mortality. The consensus explicitly recognized that although individual clinical variables carry prognostic significance, their integration into validated composite scoring systems substantially enhances predictive accuracy. Consistent with the findings of the present study, the panel further noted that the Rockall score demonstrates superior discriminative ability for mortality, whereas the GBS is better suited for identifying patients requiring urgent clinical intervention. In the 2019 update of the same consensus, the group declined to make a formal recommendation for or against the pre-endoscopic Rockall score, explicitly acknowledging the superiority of the complete Rockall score over its pre-endoscopic counterpart—a position that directly reinforces the added prognostic value of incorporating endoscopic parameters into risk assessment. Early endoscopy within 24 hours was unanimously endorsed.[8] Taken together, these evolving guideline recommendations reflect a broader trend in the literature toward a functional differentiation between the two scoring systems: the GBS is increasingly recognized as the preferred instrument for pre-endoscopic triage and safe early discharge decisions, whereas the Rockall score—by virtue of its inclusion of endoscopic findings, age, and comorbidity—is better positioned for comprehensive mortality risk stratification once endoscopic assessment has been completed. In a prospective international multicenter study by Stanley et al.,[9] five composite scoring systems—the Glasgow-Blatchford score, admission Rockall score, full Rockall score, AIMS65, and PNED score—were evaluated across 3,012 patients with UGIB. All composite systems demonstrated substantially greater predictive accuracy than individual clinical parameters, reinforcing the value of structured multivariable risk assessment. Additionally, the Glasgow-Blatchford score showed the highest discriminative ability for predicting the need for endoscopic treatment (AUROC=0.75), yet yielded the lowest predictive accuracy for 30-day mortality (AUROC=0.64), for which the AIMS65 and PNED scores performed best (AUROC=0.77 each). Because this co-

hort included both variceal and non-variceal etiologies without separate subgroup analysis, the results should be interpreted with caution specifically in the non-variceal setting. In the present study, both scoring systems demonstrated stronger associations than individual clinical and laboratory variables in predicting mortality and rebleeding, a finding consistent with the evidence outlined above and supportive of composite risk assessment over single-parameter clinical judgment.

Among the individual clinical and laboratory variables assessed in the present study, hemodynamic stability at admission demonstrated the strongest correlation with one-month mortality (r=-0.346), outperforming all other individual parameters, including age, serum urea, and hemoglobin. This finding is consistent with a 2023 systematic review and meta-analysis by Tari et al.,[10] comprising 39 cohort studies, in which hemodynamic instability at admission was identified as a strong independent predictor of adverse outcomes in acute GIB. Pooled analysis demonstrated that hemodynamically compromised patients had a 5.48-fold increased risk of in-hospital mortality (OR: 5.48; 95% CI: 3.99–7.52) and a 3.99-fold increased risk of 30-day mortality (OR: 3.99; 95% CI: 3.08–5.17), along with markedly higher rates of rebleeding and surgical intervention. Age demonstrated the second strongest correlation with one-month mortality among individual clinical parameters in the present study (r=0.332). In a prospective multicenter study by Hwang et al.[11] comprising 1,584 patients with non-variceal UGIB, multivariate analysis revealed that age >65 years independently contributed to 30-day mortality risk (OR: 2.627; 95% CI: 1.298–5.318), a finding consistent with the diminished physiological reserve and increasing comorbidity burden characteristic of older adults. Marmo et al.[12] further corroborated this finding in a prospective multicenter study of 2,317 patients, demonstrating that 85% of deaths were associated with one or more major comorbid conditions, with severe comorbidity emerging as the second strongest independent predictor of mortality after hemodynamic instability (OR: 6.72; 95% CI: 1.87–24.0).

The present study has several notable strengths. As a tertiary referral center for endoscopic intervention, the study hospital received patients from a wide catchment area, enhancing the clinical diversity and representativeness of the cohort. The strict exclusion of variceal bleeding ensured a homogeneous study population, thereby enhancing the internal validity of the findings and their specific applicability to the non-variceal setting.

The findings of this study should be interpreted in light of several notable limitations. First, this was a single-center study, which may limit the generalizability of the findings to other clinical settings and patient populations. Second, the study was conducted between 2001 and 2005, a period preceding the widespread adoption of modern endoscopic hemostatic techniques and standardized high-dose proton pump inhibitor protocols; outcomes in contemporary practice may therefore differ. Third, follow-up data on rebleeding were available only for a subset of patients because a considerable proportion of the cohort could not be reached for follow-up, limiting the statistical power of the rebleeding analyses and introducing potential selection bias. Finally, the use of Spearman correlation as the primary statistical method, although appropriate for the study objectives, does not allow identification of independent predictors after adjustment for confounding variables, thereby limiting causal inference.

CONCLUSION

In conclusion, both the Rockall and Glasgow-Blatchford scoring systems demonstrated statistically significant correlations with one-month mortality and rebleeding in non-variceal UGIB, with the Rockall score showing a stronger association with mortality. Both systems showed stronger correlations than individual clinical parameters, supporting the use of composite risk assessment in clinical practice. Among individual variables, hemodynamic stability and advanced age emerged as the strongest correlates of mortality, reinforcing the importance of early clinical assessment at the time of admission.

Disclosures

Ethics Committee Approval: The study was approved by the Istanbul University of Health Sciences Kanuni Sultan Suleyman Training and Research Hospital Scientific Research Ethics Committee (No: 3, Date: 08/01/2025).

Informed Consent: Due to the retrospective nature of the study, the requirement for written informed consent was waived by the institutional ethics committee.

Conflict of Interest Statement: The authors have no conflicts of interest to declare.

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

Use of AI for Writing Assistance: No use of AI-assisted technologies was declared by the authors.

Author Contributions: Concept – N.B.D.Ö., M.O.A.; Design – N.B.D.Ö., H.A.; Supervision – H.A.; Data collection and/or processing – N.B.D.Ö.; Data analysis and/or interpretation – C.Ö.; Literature search – M.O.A.; Writing – N.B.D.Ö., M.O.A.; Critical review – N.B.D.Ö., H.A.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Rockall TA, Logan RF, Devlin HB, Northfield TC. Risk assessment after acute upper gastrointestinal haemorrhage. Gut 1996;38:316–21. [CrossRef]
2. Blatchford O, Murray WR, Blatchford M. A risk score to predict need for treatment for upper-gastrointestinal haemorrhage. Lancet 2000;356:1318–21. [CrossRef]
3. Lu M, Sun G, Huang H, Zhang X, Xu Y, Chen S, et al. Comparison of the Glasgow-Blatchford and Rockall Scores for prediction of nonvariceal upper gastrointestinal bleeding outcomes in Chinese patients. Medicine (Baltimore) 2019;98:e15716. [CrossRef]
4. Custovic N, Husic-Selimovic A, Srsen N, Prohic D. Comparison of Glasgow-Blatchford Score and Rockall Score in Patients with Upper Gastrointestinal Bleeding. Med Arch 2020;74:270–4. [CrossRef]
5. Stanley AJ, Dalton HR, Blatchford O, Ashley D, Mowat C, Cahill A, et al. Multicentre comparison of the Glasgow Blatchford and Rockall Scores in the prediction of clinical end-points after upper gastrointestinal haemorrhage. Aliment Pharmacol Ther 2011;34:470–5. [CrossRef]
6. Kozai L, Tan A, Nebrejas K, Nishimura Y. Comparative diagnostic utility of Rockall and Glasgow-Blatchford scores in non-variceal upper gastrointestinal bleeding: a systematic review and meta-analysis. Eur J Gastroenterol Hepatol 2025;37:161–6. [CrossRef]
7. Barkun AN, Bardou M, Kuipers EJ, Sung J, Hunt RH, Martel M, et al.; International Consensus Upper Gastrointestinal Bleeding Conference Group. International consensus recommendations on the management of patients with nonvariceal upper gastrointestinal bleeding. Ann Intern Med 2010;152:101–13. [CrossRef]
8. Barkun AN, Almadi M, Kuipers EJ, Laine L, Sung J, Tse F, et al. Management of Nonvariceal Upper Gastrointestinal Bleeding: Guideline Recommendations From the International Consensus Group. Ann Intern Med 2019;171:805–22. [CrossRef]
9. Stanley AJ, Laine L, Dalton HR, Ngu JH, Schultz M, Abazi R, et al.; International Gastrointestinal Bleeding Consortium. Comparison of risk scoring systems for patients presenting with upper gastrointestinal bleeding: international multicentre prospective study. BMJ 2017;356:i6432. [CrossRef]
10. Tari E, Frim L, Stolcz T, Teutsch B, Veres DS, Hegyi P, et al. At admission hemodynamic instability is associated with increased mortality and rebleeding rate in acute gastrointestinal bleeding: a systematic review and meta-analysis. Ther Adv Gastroenterol 2023;16:17562848231190970. [CrossRef]
11. Hwang S, Jeon SW, Kwon JG, Lee DW, Ha CY, Cho KB, et al.; Daegu-Gyengbuk Gastrointestinal Study Group (DGSG). The Novel Scoring System for 30-Day Mortality in Patients with Non-variceal Upper Gastrointestinal Bleeding. Dig Dis Sci 2016;61:2002–10. [CrossRef]
12. Marmo R, Koch M, Cipolletta L, Bianco MA, Grossi E, Rotondano G; PNED 1 and PNED 2 Investigators. Predicting mortality in patients with in-hospital nonvariceal upper GI bleeding: a prospective, multicenter database study. Gastrointest Endosc 2014;79:741–9.e1. [CrossRef]

DOI: 10.14744/cm.2026.31644
Compreh Med 2026;18(Suppl. 1):11-15

RESEARCH ARTICLE

Predictive Value of Pretreatment Inflammation-Based Prognostic Scores for Primary Non-Muscle-Invasive Bladder Cancer

Haydar Güler[1], Uğur Yücetaş[2], Erkan Sönmezay[2], Erkan Erkan[2], Yaşar Pazır[3], Mustafa Kadıhasanoğlu[4]

[1]Department of Urology, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
[2]Department of Urology, İstanbul Training and Research Hospital, İstanbul, Türkiye
[3]Department of Urology, Antalya City Hospital, Antalya, Türkiye
[4]Department of Urology, University of Istanbul-Cerrahpaşa, Cerrahpaşa Faculty of Medicine, İstanbul, Türkiye

ABSTRACT

Objective: Inflammation-based prognostic scores include the neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio (PLR), and lymphocyte-to-monocyte ratio (LMR). These proposed biomarkers may serve as useful prognostic indicators for stratifying patients with favorable and poor prognoses. The aim of this study was to evaluate the predictive value of pretreatment prognostic scores for differentiating non-muscle-invasive bladder cancer (NMIBC) risk groups.

Materials and Methods: Data from patients who underwent transurethral resection of the tumor and were newly diagnosed with NMIBC between January 2015 and July 2018 were retrospectively analyzed. Patient demographics, tumor characteristics, and laboratory results were recorded. Patients were stratified into three risk groups (low-, intermediate-, and high-risk) according to the European Association of Urology guidelines. These three risk groups were compared in terms of NLR, PLR, and LMR.

Results: A total of 363 patients were included in the study. The numbers of patients in the low-, intermediate-, and high-risk groups were 79, 103, and 181, respectively. The mean NLR and LMR values were similar among the groups; however, the mean PLR value was significantly lower in the low-risk group (p=0.003). The PLR cutoff value was 148.3, and the area under the curve was 0.63, with a sensitivity of 52% and specificity of 81%.

Conclusion: This study demonstrated that preoperative PLR is a prognostic marker for distinguishing the low-risk group from the intermediate- and high-risk groups among patients with NMIBC. As an easily accessible marker, PLR can be integrated into current risk stratification models and may be useful in decision-making regarding treatment and follow-up planning.

Keywords: Bladder cancer, inflammatory index, lymphocyte, monocyte, platelet

How to cite this article: Güler H, Yücetaş U, Sönmezay E, Erkan E, Pazır Y, Kadıhasanoğlu M. Predictive Value of Pretreatment Inflammation-Based Prognostic Scores for Primary Non-Muscle-Invasive Bladder Cancer. Compreh Med 2026;18(Suppl. 1):11-15.

INTRODUCTION

Bladder cancer (BC), the most common malignancy of the urinary tract, is the ninth most common cancer worldwide.[1] Global data from 2020 indicate that this malignancy is associated with a high disease burden, with 573,278 new cases and 212,536 deaths annually.[2] Urothelial carcinoma is the predominant histological subtype of bladder cancer, with non-muscle-invasive bladder cancer (NMIBC) accounting for nearly 75–85% of cases at the time of clinical presentation.[3] For these patients, transurethral resection of the bladder tumor (TUR-BT) remains the gold-standard primary intervention. In addition, adjuvant intravesical treatments are recommended according to patients' risk classifications. In NMIBC, recurrence and progression rates can reach 70% and 30%, respectively, and vary widely among patients.[4] Therefore, risk tables have been developed to accurately predict

Address for Correspondence: Haydar Güler, Department of Urology, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
E-mail: drguler44@gmail.com **ORCID ID:** 0000-0002-7167-5449

Received date: 13.05.2026
Revised date: 12.06.2026
Accepted date: 13.06.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

recurrence and progression and to determine individualized treatment and follow-up strategies.

The European Organisation for Research and Treatment of Cancer (EORTC) and Spanish Urological Club for Oncological Treatment (CUETO) risk tables serve as essential prognostic tools, allowing estimation of recurrence and progression probabilities based on established scoring criteria.[4,5] However, the predictive accuracy of these models is not optimal, and additional predictive markers are needed.[6]

Current evidence indicates that cancer can induce a systemic inflammatory response, resulting in alterations in circulating immune cells. The most notable changes are neutrophilia accompanied by relative lymphocytopenia or thrombocytosis. These immune cells contribute to tumor cell proliferation, differentiation, and maturation within the tumor microenvironment by releasing various cytokines and chemokines.[7] There is a well-established correlation between oncologic prognosis and various inflammation-based scores, such as the preoperative platelet-to-lymphocyte ratio (PLR), neutrophil-to-lymphocyte ratio (NLR), and lymphocyte-to-monocyte ratio (LMR). The predictive utility of these preoperative ratios has been validated across a broad spectrum of malignancies, including breast, lung, liver, and colon cancers.[8–10] In recent years, the prognostic value of inflammation-based scores in BC has been investigated in numerous studies, with varying results.[11–14] However, there is insufficient evidence regarding whether these biomarkers have prognostic value in differentiating patients with NMIBC according to risk groups.

This study aimed to assess the predictive utility of preoperative inflammation-based prognostic scores in stratifying patients with NMIBC according to risk groups.

MATERIALS and METHODS

Patients

Data from patients who underwent TUR-BT and were newly diagnosed with NMIBC at two different institutions between January 2015 and July 2018 were retrospectively analyzed. Patients with acute infection, autoimmune disease, hepatic or renal dysfunction, hematological malignancy, or other malignant diseases were excluded from the study.

Patient demographics, tumor characteristics, and laboratory results were recorded. Tumor stage was determined according to the 2002 TNM classification, and tumor grade was determined according to the 2004 World Health Organization classification system.[15,16] Based on the criteria established in the European Association of Urology (EAU) guidelines, patients were categorized into low-, intermediate-, and high-risk groups.[3] These three risk groups were compared in terms of NLR, PLR, and LMR.

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki and was approved by University of Health Sciences İstanbul Training and Research Hospital Clinical Research Ethics Committee (Approval Number: 94, Date: 27.04.2026).

Blood Analysis

As part of the standard preoperative assessment, venous blood samples were obtained from all participants. These samples, collected in tubes containing ethylenediaminetetraacetic acid, underwent hematological evaluation using Sysmex XE-2100 and XE-5000 analyzers (Sysmex UK, Milton Keynes, UK). Complete blood count results obtained within 60 days before surgery were evaluated. For patients with multiple measurements, the values from the sample obtained closest to the date of surgery were used. Neutrophil, lymphocyte, monocyte, and platelet counts were obtained, and the NLR, PLR, and LMR values for each patient were calculated. NLR and PLR were calculated by dividing the neutrophil and platelet counts by the lymphocyte count, respectively. LMR was calculated by dividing the lymphocyte count by the monocyte count.

Statistical Analysis

Data were presented as means, medians (IQRs), and ranges. Data normality was assessed using the Kolmogorov-Smirnov test. For subsequent statistical analyses, the chi-square and Kruskal-Wallis tests were used, with $p<0.05$ considered statistically significant. Additionally, the optimal PLR cutoff value was determined using the Youden index and receiver operating characteristic (ROC) analysis. Statistical analyses were performed using STATA version 11.0 (StataCorp LP, College Station, TX, USA).

RESULTS

A total of 363 patients (315 men and 48 women) were evaluated, with a mean age of 64.25±10.54 years. In terms of risk distribution, the high-risk group was the largest (n=181), whereas the intermediate- and low-risk groups included 103 and 79 patients, respectively. The mean age of the low-risk group was lower than those of the intermediate- and high-risk groups (p=0.0001). Sex distribution and smoking status were similar among the groups (Table 1). The mean NLR and LMR values were also similar among the groups; however, the mean PLR value was significantly lower in the low-risk group

Table 1. Comparison of patient characteristics and pretreatment inflammation-based prognostic scores among the risk groups

Variable	Low-risk group (n=79)	Intermediate-risk group (n=103)	High-risk group (n=181)	p
Age (years), mean±SD	58.3±10.1	64.5±9.9	66.6±9.5	0.0001*
Sex (male/female)	70/9	88/14	156/24	0.9
Smoking status (smokers/non-smokers)	50/7	68/7	123/14	0.8
NLR, median (IQR)	2.26 (1.46)	2.28 (1.48)	2.3 (1.9)	0.6
LMR, median (IQR)	3.89 (2.13)	3.79 (1.78)	3.48 (2.28)	0.7
PLR, median (IQR)	102 (49)	118 (63)	124 (69)	0.003*

*: p<0.05, statistically significant. IQR: interquartile range; LMR: lymphocyte-to-monocyte ratio; NLR: neutrophil-to-lymphocyte ratio; PLR: platelet-to-lymphocyte ratio; SD: standard deviation.

(p=0.003) (Table 1). For predicting the low-risk group, the PLR cutoff value was 148.3, and the area under the curve was 0.63, with a sensitivity of 52% and specificity of 81% (Fig. 1).

DISCUSSION

In the present study, we evaluated the value of three pretreatment inflammation-based prognostic scores (PLR, NLR, and LMR) in predicting NMIBC risk groups. Our results revealed that PLR was significantly lower in patients with low-risk NMIBC than in those in the intermediate- and high-risk groups. Conversely, neither NLR nor LMR showed significant differences among the risk categories in our study population. To the best of our knowledge, the current literature lacks prior research specifically exploring the association between PLR and risk group classification in NMIBC.

The clinical management of bladder cancer is complex because of its inherent biological heterogeneity, requiring precise staging and risk assessment to tailor therapeutic interventions. Although histopathological examination remains the definitive diagnostic tool, staging discrepancies are common; evidence suggests that nearly half of patients undergoing radical cystectomy may have been incorrectly staged preoperatively.[17,18] In recent years, efforts have focused on improving staging accuracy by integrating histopathological findings with imaging modalities, such as computed tomography and magnetic resonance imaging, as well as molecular biomarkers into predictive nomograms.[19,20] Although several such models incorporating molecular parameters have been developed, there remains a need to include additional variables, particularly laboratory-based markers, to further improve their predictive performance.

Accumulating evidence indicates that the host inflammatory response plays a key role in carcinogenesis, as well as in tumor growth and progression.[7,21] Neutrophils and platelets, which are important components of the immune system, have inhibitory effects on antitumor activity. In contrast, lymphocytes have an activating effect. Lymphocytes contribute to the host's antitumor immunity by mediating cytotoxic cell death and inhibiting tumor cell proliferation and metastasis. Reduced lymphocyte counts weaken the host's antitumor immune response, thereby increasing the likelihood of cancer recurrence and progression.[22,23] Conversely, platelets facilitate tumor angiogenesis by releasing proangiogenic factors and also contribute to protecting tumor cells from cytolytic destruction.[24,25] Monocytes can increase the number of tumor-associated macrophages, which may contribute to tumor growth and metastasis and therefore reflect tumor burden.[26] Additionally, it has been suggested that neutrophilia promotes tumor growth and progression by increasing inflammatory mediators such as proangiogenic factors (VEGF), proteases, growth factors (CXCL8), and antiapoptotic markers (NF-κB).[27] Therefore,

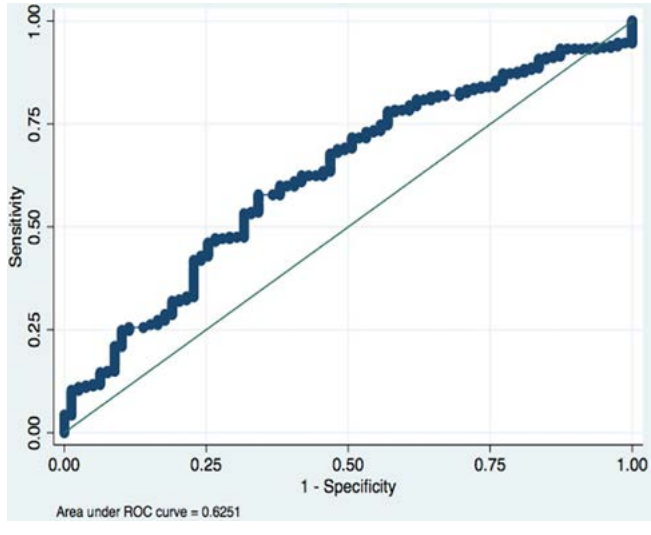


Figure 1. ROC curve of platelet-lymphocyte ratio to predict low-risk group

ROC: Receiver operating characteristic

the numbers of these immune cells and their ratios to one another may reflect the host immune response and serve as predictive markers of cancer prognosis.

In the present study, our findings showed that preoperative PLR has predictive value in distinguishing the low-risk NMIBC group from the intermediate- and high-risk groups. Similarly, previous studies involving patients with NMIBC reported that higher preoperative PLR was associated with higher tumor grade and lamina propria invasion.[28,29] Supporting our findings, tumor grade and lamina propria invasion are among the parameters used to stratify patients with NMIBC into risk groups. Furthermore, Wang et al.'s[30] meta-analysis of 3,303 patients with BC revealed that lower PLR was associated with better overall survival. Taken together, PLR may be considered a low-cost prognostic marker for distinguishing patients with NMIBC according to risk group and can be easily assessed before surgery.

The majority of studies investigating the prognostic significance of preoperative inflammatory markers in NMIBC and MIBC have shown that high NLR and low LMR are associated with worse oncologic outcomes. Several studies have reported that high NLR levels are associated with muscle-invasive disease at TUR-BT. Furthermore, Viers et al.[31] reported an increased risk of extravesical disease extension and lymph node metastasis in patients with higher NLR. In their meta-analysis, Tang et al.[32] included studies involving both NMIBC and muscle-invasive bladder cancer (MIBC) and concluded that high NLR was associated with poorer overall survival (OS), cancer-specific survival (CSS), recurrence-free survival (RFS), and progression-free survival. Regarding LMR, Adamkiewicz et al.[33] reported that LMR was an independent predictor of progression in patients with NMIBC receiving intravesical BCG therapy. Moreover, Ma et al.'s[34] meta-analysis revealed that lower LMR was associated with worse OS, CSS, and RFS in patients with MIBC. However, in the present study, preoperative NLR and LMR had no predictive value in distinguishing patients with NMIBC according to risk group. In addition to its statistical significance, the clinical applicability of PLR deserves consideration. Since a complete blood count is routinely obtained before TUR-BT, PLR can be calculated without additional cost or patient burden. Therefore, PLR may serve as a practical adjunctive marker to support current risk stratification systems, particularly in situations in which pathological risk assessment is uncertain or additional prognostic information is desired.

The limitations of this study include its retrospective design and comparatively small sample size. Furthermore, because survival data were not available, we could not analyze the relationship between inflammation-based prognostic scores and long-term oncological outcomes such as recurrence, progression, cancer-specific survival, or overall survival. Another limitation is that inflammatory markers were assessed using a single preoperative blood sample, which may not fully reflect dynamic changes in systemic inflammatory status over time. In addition, although patients with active inflammatory conditions were excluded, residual confounding factors affecting blood cell counts cannot be completely ruled out. Therefore, prospective multicenter studies with larger cohorts and long-term follow-up are warranted to validate our findings.

CONCLUSION

This study demonstrated that preoperative PLR is a prognostic marker for distinguishing the low-risk group from the intermediate- and high-risk groups among patients with NMIBC. As an easily accessible and inexpensive biomarker, PLR may complement existing clinicopathological risk stratification systems and contribute to individualized treatment and surveillance strategies. Nevertheless, the relatively modest discriminative performance observed in the present study indicates that PLR should be considered an adjunct rather than a standalone predictor. Further prospective studies are needed before its routine incorporation into clinical practice.

Disclosures

Ethics Committee Approval: The study was approved by the University of Health Sciences İstanbul Training and Research Hospital Clinical Research Ethics Committee (No: 94, Date: 27/04/2026).

Informed Consent: The requirement for informed consent was waived by the ethics committee due to the retrospective nature of the study.

Conflict of Interest Statement: The authors have no conflicts of interest to declare.

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

Use of AI for Writing Assistance: No use of AI-assisted technologies was declared by the authors.

Author Contributions: Concept – H.G.; Design – Y.P.; Supervision – M.K.; Funding – E.S., E.E.; Materials – U.Y., Y.P.; Data collection and/or processing – H.G.; Data analysis and/or interpretation – H.G., E.E.; Literature search – U.Y., E.S.; Writing – H.G., Y.P.; Critical review – M.K.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin 2014;64:9–29. Erratum in: CA Cancer J Clin 2014;64:364. [CrossRef]
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J Clin 2021;71:209–49. [CrossRef]
3. Babjuk M, Burger M, Capoun O, Cohen D, Compérat EM, Dominguez Escrig JL, et al. European Association of Urology Guidelines on Non-muscle-invasive Bladder Cancer (Ta, T1, and Carcinoma in Situ). Eur Urol 2022;81:75–94. [CrossRef]
4. Sylvester RJ, van der Meijden AP, Oosterlinck W, Witjes JA, Bouffioux C, Denis L, et al. Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. Eur Urol 2006;49:466–5; discussion 475–7. [CrossRef]
5. Fernandez-Gomez J, Madero R, Solsona E, Unda M, Martinez-Piñeiro L, Gonzalez M, et al. Predicting nonmuscle invasive bladder cancer recurrence and progression in patients treated with bacillus Calmette-Guerin: the CUETO scoring model. J Urol 2009;182:2195–203. [CrossRef]
6. Soukup V, Čapoun O, Cohen D, Hernández V, Burger M, Compérat E, et al. Risk Stratification Tools and Prognostic Models in Non-muscle-invasive Bladder Cancer: A Critical Assessment from the European Association of Urology Non-muscle-invasive Bladder Cancer Guidelines Panel. Eur Urol Focus 2020;6:479–89. [CrossRef]
7. Coussens LM, Werb Z. Inflammation and cancer. Nature 2002;420:860–7. [CrossRef]
8. Kumar R, Geuna E, Michalarea V, Guardascione M, Naumann U, Lorente D, et al. The neutrophil-lymphocyte ratio and its utilisation for the management of cancer patients in early clinical trials. Br J Cancer 2015;112:1157–65. [CrossRef]
9. Roxburgh CS, McMillan DC. Role of systemic inflammatory response in predicting survival in patients with primary operable cancer. Future Oncol 2010;6:149–63. [CrossRef]
10. Zhou X, Du Y, Huang Z, Xu J, Qiu T, Wang J, et al. Prognostic value of PLR in various cancers: a meta-analysis. PLoS One 2014;9:e101119. [CrossRef]
11. Mano R, Baniel J, Shoshany O, Margel D, Bar-On T, Nativ O, et al. Neutrophil-to-lymphocyte ratio predicts progression and recurrence of non-muscle-invasive bladder cancer. Urol Oncol 2015;33:67.e1–7. [CrossRef]
12. D'Andrea D, Moschini M, Gust K, Abufaraj M, Özsoy M, Mathieu R, et al. Prognostic Role of Neutrophil-to-Lymphocyte Ratio in Primary Non-muscle-invasive Bladder Cancer. Clin Genitourin Cancer 2017;15:e755–64. [CrossRef]
13. Krane LS, Richards KA, Kader AK, Davis R, Balaji KC, Hemal AK. Preoperative neutrophil/lymphocyte ratio predicts overall survival and extravesical disease in patients undergoing radical cystectomy. J Endourol 2013;27:1046–50. [CrossRef]
14. Ceylan C, Doluoglu OG, Keleş I, Gazel E, Temuçin T, Odabaş Ö, et al. Importance of the neutrophil-to-lymphocyte ratio in muscle-invasive and non-muscle invasive bladder tumors. Urologia 2014;81:120–4. [CrossRef]
15. Montironi R, Lopez-Beltran A. The 2004 WHO classification of bladder tumors: a summary and commentary. Int J Surg Pathol 2005;13:143–53. [CrossRef]
16. Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17:1471–4. [CrossRef]
17. Shariat SF, Palapattu GS, Karakiewicz PI, Rogers CG, Vazina A, Bastian PJ, et al. Discrepancy between clinical and pathologic stage: impact on prognosis after radical cystectomy. Eur Urol 2007;51:137–49; discussion 149–51. [CrossRef]
18. Ark JT, Keegan KA, Barocas DA, Morgan TM, Resnick MJ, You C, et al. Incidence and predictors of understaging in patients with clinical T1 urothelial carcinoma undergoing radical cystectomy. BJU Int 2014;113:894–9. [CrossRef]
19. Karakiewicz PI, Shariat SF, Palapattu GS, Gilad AE, Lotan Y, Rogers CG, et al. Nomogram for predicting disease recurrence after radical cystectomy for transitional cell carcinoma of the bladder. J Urol 2006;176:1354–61; discussion 1361–2. [CrossRef]
20. Margel D, Harel A, Yossepowitch O, Baniel J. A novel algorithm to improve pathologic stage prediction of clinically organ-confined muscle-invasive bladder cancer. Cancer 2009;115:1459–64. [CrossRef]
21. Mantovani A, Allavena P, Sica A, Balkwill F. Cancer-related inflammation. Nature 2008;454:436–44. [CrossRef]
22. Pfeffer CM, Singh ATK. Apoptosis: A Target for Anticancer Therapy. Int J Mol Sci 2018;19:448. [CrossRef]
23. Zou Y, Xie J, Zheng S, Liu W, Tang Y, Tian W, et al. Leveraging diverse cell-death patterns to predict the prognosis and drug sensitivity of triple-negative breast cancer patients after surgery. Int J Surg 2022;107:106936. [CrossRef]
24. Sierko E, Wojtukiewicz MZ. Platelets and angiogenesis in malignancy. Semin Thromb Hemost 2004;30:95–108. [CrossRef]
25. Sierko E, Wojtukiewicz MZ. Inhibition of platelet function: does it offer a chance of better cancer progression control? Semin Thromb Hemost 2007;33:712–21. [CrossRef]
26. Galdiero MR, Bonavita E, Barajon I, Garlanda C, Mantovani A, Jaillon S. Tumor associated macrophages and neutrophils in cancer. Immunobiology 2013;218:1402–10. [CrossRef]
27. Paramanathan A, Saxena A, Morris DL. A systematic review and meta-analysis on the impact of pre-operative neutrophil lymphocyte ratio on long term outcomes after curative intent resection of solid tumours. Surg Oncol 2014;23:31–9. [CrossRef]
28. Chen H, Wu X, Wen Z, Zhu Y, Liao L, Yang J. The Clinicopathological and Prognostic Value of NLR, PLR and MLR in Non-Muscular Invasive Bladder Cancer. Arch Esp Urol 2022;75:467-71. [CrossRef]
29. Atik YT, Cimen HI, Gul D, Arslan S, Kose O, Halis F. Are the preoperative neutrophil/lymphocyte ratio and platelet/lymphocyte ratio predictive for lamina propria invasion in aging patients? Aging Male 2020;23:1528–32. [CrossRef]
30. Wang X, Ni X, Tang G. Prognostic Role of Platelet-to-Lymphocyte Ratio in Patients With Bladder Cancer: A Meta-Analysis. Front Oncol 2019;9:757. [CrossRef]
31. Viers BR, Boorjian SA, Frank I, Tarrell RF, Thapa P, Karnes RJ, et al. Pretreatment neutrophil-to-lymphocyte ratio is associated with advanced pathologic tumor stage and increased cancer-specific mortality among patients with urothelial carcinoma of the bladder undergoing radical cystectomy. Eur Urol 2014;66:1157–64. [CrossRef]
32. Tang X, Du P, Yang Y. The clinical use of neutrophil-to-lymphocyte ratio in bladder cancer patients: a systematic review and meta-analysis. Int J Clin Oncol 2017;22:817–25. [CrossRef]
33. Adamkiewicz M, Bryniarski P, Kowalik M, Burzyński B, Rajwa P, Paradysz A. Lymphocyte-to-Monocyte Ratio Is the Independent Prognostic Marker of Progression in Patients Undergoing BCG-Immunotherapy for Bladder Cancer. Front Oncol 2021;11:655000. [CrossRef]
34. Ma JY, Hu G, Liu Q. Prognostic Significance of the Lymphocyte-to-Monocyte Ratio in Bladder Cancer Undergoing Radical Cystectomy: A Meta-Analysis of 5638 Individuals. Dis Markers 2019;2019:7593560. [CrossRef]

DOI: 10.14744/cm.2026.02360
Compreh Med 2026;18(Suppl. 1):16-24

RESEARCH ARTICLE

The Effects of Mild and Moderate Hypothermia on Liver Regeneration, Liver Damage, and Apoptosis in Hepatectomized Rats

Ali Önder Devay[1], Mehmet Özdoğan[2], Pamir Eren Ersoy[3], Rıza Haldun Gündoğdu[3], Recep Aktimur[4], Selma Göğkuş Beyeç[5], Seda Duygulu Devay[6], Neşe Lortlar Ünlü[7]

[1]Department of General Surgery, Medicana Bursa Hospital, Bursa, Türkiye
[2]Department of General Surgery, Private Practice, Adana, Türkiye
[3]Department of General Surgery, Yozgat Bozok University Faculty of Medicine, Yozgat, Türkiye
[4]Department of General Surgery, Samsun Liv Hospital, Samsun, Türkiye
[5]Department of General Surgery, Bursa City Hospital, Bursa, Türkiye
[6]Department of Medical Biochemistry, Bursa City Hospital, Bursa, Türkiye
[7]Department of Histology and Embryology, Atlas University Faculty of Medicine, Istanbul, Türkiye

ABSTRACT

Objective: The aim of this study was to investigate the effects of mild and moderate hypothermia on liver damage, apoptosis, and regenerative capacity after hepatectomy and the mechanisms underlying these effects.

Materials and Methods: A total of 34 male Wistar Albino rats were divided into 3 groups. The core temperatures of 10 rats in the normothermia group (Group 1) were maintained at 37°C using an external lamp. In the hypothermia groups, an external ice-cooling method was used. The core temperatures of 12 rats in the mild hypothermia group (Group 2) were reduced to 32–34°C, and those of 12 rats in the moderate hypothermia group (Group 3) were reduced to 26–28°C. Approximately 76% hepatectomy was performed in all rats, and the rats were followed for 48 hours. MDA and GSH (glutathione) levels were measured in liver tissue, and AST, ALT, GSH, and MDA levels were measured in blood samples. Regenerated liver weight was calculated, and liver tissue was evaluated histopathologically.

Results: Tissue GSH levels decreased in Group 2 compared with Group 1, whereas serum MDA levels increased. Tissue GSH levels were significantly higher in Group 3 than in Group 1. There was no difference among the groups in the liver damage index. There was a significant increase in PCNA (proliferating cell nuclear antigen) activity in Groups 2 and 3 compared with Group 1. There was a significant decrease in caspase-9 activity in Groups 2 and 3 compared with Group 1. There was no difference among the groups in regenerated liver weight measurements.

Conclusion: Systemic mild hypothermia applied during hepatectomy increased oxidative stress, whereas moderate hypothermia caused oxidative stress levels to return to normal. Mild and moderate hypothermia had a positive effect by increasing regenerative capacity and suppressing apoptosis after liver resection.

Keywords: Apoptosis, hepatectomy, hypothermia, liver damage, liver regeneration

How to cite this article: Devay AÖ, Özdoğan M, Ersoy PE, Gündoğdu RH, Aktimur R, Göğkuş Beyeç S, et al. The Effects of Mild and Moderate Hypothermia on Liver Regeneration, Liver Damage, and Apoptosis in Hepatectomized Rats. Compreh Med 2026;18(Suppl. 1):16-24.

INTRODUCTION

Stability of liver function is essential for body homeostasis; for this purpose, multiple regenerative mechanisms have evolved to ensure that adequate liver function is always maintained.[1] In the biological context, the liver is the only internal organ that has the ability to regenerate after partial resection and chemical injury. This property is technically defined as hyperplasia of the remaining hepatocytes to compensate for lost tissue, restore full liver size, and meet the metabolic needs of the organism without recovery.[2]

Address for Correspondence: Ali Önder Devay, Department of General Surgery, Medicana Bursa Hospital, Bursa, Türkiye
E-mail: aonderdevay@gmail.com **ORCID ID:** 0000-0003-3430-0215

Received date: 15.05.2026
Revised date: 18.06.2026
Accepted date: 23.06.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Ischemia/reperfusion (I/R) injury is mainly caused by the continuous metabolic demands of a warm ischemic organ and the lack of oxygen to meet these demands. One way to increase the ischemic resistance of the liver is to lower the organ temperature.[3] Hypothermia causes morphological and circulatory changes in the liver. Systemic hypothermia may have a negative effect on liver tissue by causing cellular changes such as a decreased sinusoid-to-hepatocyte volume ratio and *ex vivo* hepatocyte necrosis.[4] On the other hand, studies have also shown that hypothermia has a positive effect on liver damage.[5,6] Experimental studies have demonstrated that mild hypothermia (34°C) decreased liver damage and increased 24-hour survival in rats with I/R injury.[7] Mild external hypothermia (32–35°C) has been shown to decrease plasma aspartate transaminase (AST) levels and histological damage, alleviate apoptosis, and increase survival in mice with acetaminophen (APAP)-induced liver injury.[6]

The aim of this study was to investigate the effects of mild and moderate hypothermia on liver damage, apoptosis, and regenerative capacity after hepatectomy and to compare these effects.

MATERIALS and METHODS

Experimental Protocol

This study was approved by the Ethics Committee of Başkent University (29.04.2008/2008/AP-455). The study was conducted in compliance with the Declaration of Helsinki and local legal regulations. In this study, artificial intelligence tools were used for language editing purposes and to reduce the reference list to 250 words.

A total of 34 adult male Wistar Albino rats weighing 275–375 g were used in the study. Since mortality was observed in 5 of the 34 rats (1 in Group 1, 2 in Group 2, and 2 in Group 3), the study was completed with 29 rats. The animals were housed under a 12-hour light-dark cycle and were given tap water and a standard laboratory rat diet.

The rats were divided into 3 groups (Group 1, n=10; Groups 2 and 3, n=12). Group 1 was the normothermia group. The body temperature of the rats in this group was maintained at 36–37°C using an external heating lamp. Group 2 was the mild hypothermia group. In this group, body temperature was reduced to 32–34°C by spontaneous temperature reduction, and core temperature was kept constant by monitoring without intervention or by applying external ice molds. Group 3 was the moderate hypothermia group. In this group, core temperature was reduced to 26–28°C using an external cooling method and was kept constant using ice molds. Body temperature was monitored using a rectal probe throughout the procedure.

The study was conducted in three phases. In the first phase, the skin of all animals was shaved and sterilized with 10% povidone-iodine. General anesthesia with ketamine and xylazine was administered, and laparotomy was performed under sterile conditions. The desired core temperatures were reached for each group by monitoring rectal temperature and were maintained for 30 minutes. In the second phase, partial hepatectomy involving the caudate lobe, left lateral lobe, and median lobe was performed. The weight of the resected liver tissue was measured. The animals were resuscitated intraperitoneally with 2 mL of normal saline, and the incision was closed. In the third phase, the rats were kept under observation for 48 hours at room temperature. During this period, fentanyl (0.02 mg/kg, SC, twice daily) was administered as an analgesic.

Forty-eight hours postoperatively, the rats were anesthetized with ketamine (200 mg/kg, intramuscularly) and sacrificed by intracardiac blood collection. The abdomen was opened through the previous incision, and the liver tissues were removed and weighed using a precision balance. The tissues were washed with 0.9% NaCl and divided into two equal parts. Half of the samples were wrapped in aluminum foil, immersed in liquid nitrogen for biochemical examinations, and stored at -80°C until analysis. The remaining tissues were fixed in 10% formalin solution for histopathological examination, processed using standard histological tissue-processing methods, and embedded in paraffin blocks. Blood samples were centrifuged at 4000×g for 10 minutes. The obtained plasma was placed in Eppendorf tubes for AST (aspartate aminotransferase) and ALT (alanine aminotransferase) measurements, and the serum was placed in Eppendorf tubes for MDA (malondialdehyde) and GSH (glutathione) determination and stored at -80°C until analysis.

Biochemical Examinations

Homogenization of Liver Tissue Samples

Liver tissue samples were removed from the deep freezer when the biochemical parameters MDA and GSH were to be analyzed. A total of 100 mg of tissue was mixed with saline solution at a 1:10 ratio (w/v) and homogenized in an ice-water bath at 4000 rpm for 1–1.5 minutes. The crude extract (homogenate) was centrifuged in a refrigerated centrifuge at 12,000 rpm for 20 minutes. The resulting supernatants were separated for MDA and GSH determination.

Determination of MDA in Rat Liver Tissue Samples

For MDA determination, 50 µL of tissue supernatant was collected, and 50 µL of 8.1% SDS was added. Then, 375 µL of 0.8% TBA, 375 µL of 20% acetic acid (pH=3.5), and 150 µL of distilled water were added. The resulting mixture was boiled in a water bath at 95°C for 60 minutes. Immediately afterward, the samples were cooled, and 250 µL of distilled water and 1,250 mL of n-butanol/pyridine (15:1, v/v) were added and vortexed. The samples were then centrifuged at 4000 rpm for 10 minutes. A standard curve was obtained using 1,1,3,3-tetramethoxypropane as a standard for MDA, and the amount of MDA in nmol was determined from the absorbance values of the samples. The results were expressed as nmol/g wet tissue.[8]

Serum MDA Determination

For serum MDA determination, 250 µL of serum samples were collected, 1.25 mL of 20% TCA and 0.5 mL of 0.67% TBA were added, and the samples were boiled in a hot-water bath for 30 minutes. They were then immediately cooled, 2 mL of n-butanol was added to all tubes, and the samples were centrifuged at 3000 rpm for 5 minutes. The supernatant was read against n-butanol at 532 nm.

Determination of GSH in Rat Liver Tissue Samples and Serum

Tissue and serum GSH levels were determined by calculating total thiol according to the method described by Ohkawa et al.[7]

For serum total thiol determination, 250 µL of serum samples were collected, and 750 µL of TRIS-EDTA (pH=8.2), 50 µL of DTNB, and 3950 µL of pure methanol were added. To the blank tubes, 250 µL of distilled water, 750 µL of TRIS-EDTA (pH=8.2), 50 µL of DTNB, and 3950 µL of pure methanol were added. The tubes containing the samples and blanks were left in the dark for 15 minutes and vortexed every 5 minutes. At the end of the 15-minute period, the tubes were centrifuged at 3000×g, and the samples were read against a blank at 412 nm. For tissue total thiol determination, 250 µL of the supernatant samples prepared as described above were collected, and 750 µL of TRIS-EDTA (pH=8.2), 50 µL of DTNB, and 3950 µL of pure methanol were added. To the blank tubes, 250 µL of distilled water, 750 µL of TRIS-EDTA (pH=8.2), 50 µL of DTNB, and 3950 µL of pure methanol were added. The tubes containing the samples and blanks were left in the dark for 15 minutes and vortexed every 5 minutes. At the end of the 15-minute period, the tubes were centrifuged at 3000×g, and the samples were read against a blank at 412 nm.

Histochemical Examinations

Tissue samples obtained from the rats were washed with saline solution, fixed in 10% formalin solution for 72 hours, and embedded in paraffin blocks. Sections 4–5 µm thick were placed on slides, deparaffinized first in an oven and then in a series of xylene solutions, and subsequently rehydrated through a graded alcohol series. The sections were stained with hematoxylin and eosin and evaluated under a light microscope.

To determine liver damage, tissues stained with hematoxylin and eosin were evaluated for degeneration, sinusoidal dilation, sinusoidal congestion, and inflammatory cell infiltration.

The following scoring system was used for the samples, which were evaluated separately for each criterion:

0: No findings

1: Findings in less than 25% of the area

2: Findings in 25–50% of the area

3: Findings in more than 50% of the cross-sectional area.[9]

Immunohistochemical Method

The fixed tissues were processed using conventional light microscopy methods, and paraffin blocks were prepared. Sections 4–5 µm thick were obtained from the prepared paraffin blocks and mounted on polylysine-coated slides. After deparaffinization, the slides were incubated with PCNA (Cat#: MS-106-P, Thermo, Fremont, CA, USA) primary antibody diluted 1:200 and caspase-9 (Cat#: RB-1205-P, Thermo, Fremont, CA, USA) primary antibody diluted 1:100 for 60 minutes at room temperature. After washing the slides with PBS, a biotin-containing secondary antibody (Thermo, USA) was applied for 20 minutes to allow binding to the primary antibody. The slides were then washed with PBS and incubated with the streptavidin-peroxidase enzyme complex (Thermo, USA) for 20 minutes, allowing the enzyme complex to bind to biotin on the slides.

Finally, the chromogen AEC (3-amino-9-ethylcarbazole) (TA-125-HA, Thermo, Fremont, CA, USA) was added and incubated for approximately 10 minutes, allowing the visible reaction product to develop.

Mayer's hematoxylin (0B524092, Merck) was used as the background stain. The stained slides were covered with coverslips using Ultramount and evaluated under a computer-equipped photomicroscope (DM 4000B, Leica, Germany).

To evaluate PCNA, positively stained hepatocytes were counted in 5 randomly selected areas at ×400 magnification.[10]

Table 1. Statistical evaluation of GSH and MDA levels in liver tissue and serum and serum AST and ALT measurements (mean±SEM)

	MDA		GSH		AST (U/L)	ALT (U/L)
	SERUM (nmol/ml)	TISSUE (nmol/gr tissue)	SERUM (µmol/l)	TISSUE (µmol/gr tissue)		
Group 1	8.35±0.26	113.20±21.93	253.77±17.02	17.75±1.13	416.67±48.73	291.00±48.51
Group 2	11.76±0.78[b]	96.24±11.11	260.61±34.19	13.22±0.84[a]	655.20±191.39	466.30±151.29
Group 3	9.27±1.29	115.90±19.26	359.54±65.74	18.50±1.10[c]	683.50±166.28	526.80±154.43

[a]: p: 0.01, mild hypothermia decreased tissue GSH levels; [b]: p: 0.01, serum MDA levels were higher in mild hypothermia group; [c]: p: 0.01, in moderate hypothermia group, tissue GSH levels were increased compared to mild hypothermia group. ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; GSH: Glutathione; MDA: Malondialdehyde; SEM: Standard error of the mean; U/L: Units per liter; nmol/mL: Nanomoles per milliliter; nmol/g tissue: Nanomoles per gram of tissue; µmol/L: Micromoles per liter; µmol/g tissue: Micromoles per gram of tissue

Table 2. Liver damage index

	Degeneration	Sinozoidal dilatation	Sinozoidal congestion	Cell infiltration	Total score
Group 1	2.22±0.32	2.22±0.28	0.20±0.13	2.22±0.32	5.44±0.65
Group 2	2.10±0.18	2.40±0.22	0.36±0.12	2.22±0.32	6.10±0.38
Group 3	1.60±0.16	1.80±0.20	0.25±0.32	2.22±0.32	5.40±0.30

For caspase-9 evaluation, the entire preparation was scanned, and cells showing positive staining were assessed as follows:

0: No staining

1: Mild staining

2: Moderate staining

3: Strong staining.[11]

Statistical Analysis

Statistical analyses were performed using SPSS version 16.0 (SPSS Inc., Chicago, IL, USA). The normality of data distribution was assessed using the Shapiro–Wilk test. Due to the small sample sizes and non-normal distribution of several parameters, nonparametric tests were employed. Comparisons among the three groups were performed using the Kruskal–Wallis H test, followed by post hoc pairwise comparisons using the Mann–Whitney U test. A Bonferroni correction was applied for multiple comparisons, and the adjusted significance level was set at $p<0.017$. All data are presented as mean±standard error of the mean (SEM). A p-value<0.05 was considered statistically significant.

RESULTS

The biochemical findings and liver damage index were analyzed (Tables 1, 2). According to the biochemical findings, no statistically significant differences were found in AST or ALT levels among the groups. There was a statistically significant difference among the groups in tissue GSH levels. When the mild hypothermia group (Group 2) was compared with the normothermia group (Group 1), a statistically significant decrease in tissue GSH levels was observed. Serum MDA levels were significantly higher in the normothermia group (Group 1). When the moderate hypothermia group (Group 3) was compared with the mild hypothermia group (Group 2), a significant increase in tissue GSH levels was observed. When the immunohistochemical findings were analyzed (Table 3), statistically significant differences were found among the groups in terms of caspase-9 and PCNA activity. The following formula was used to calculate regenerated liver weight. No significant difference was found among the groups in terms of regenerated liver weight (Table 4).

Table 3. Immunohistochemical comparison of groups (Mean±SEM)

	CASPASE	PCNA
Group 1	1.51±0.17	44.66±1.13
Group 2	0.64±0.13[a]	59.70±1.41[b]
Group 3	0.79±0.10[a]	68.87±1.24[b,c]

[a]: p=0.01, mild and moderate hypothermia decreases caspase enzyme activity compared to normothermia; [b]: p=0.01, mild and moderate hypothermia increases PCNA number compared to normothermia; [c]: p=0.01, moderate hypothermia increases the number of PCNAs compared to mild hypothermia. PCNA: Proliferating cell nuclear antigen; SEM: Standard error of the mean; p: Probability value (p-value)

Regenerated liver weight=weight of residual lobes/body weight×100

Table 4. Morphological evaluation of liver according to experimental groups

	Rat weight (gr)	Rat liver weight (gr)	Resected liver weight (gr)	Liver weight at autopsy (gr)	Regenerated liver weight (gr)
Group A					
1	250	8.5	6.6	4.9	35.3
2	358	12.2	8.1	5.9	14.7
3	332	11.3	7.4	7.0	27.4
4	220	7.4	5.8	5.5	52.7
5	376	12.8	7.8	7.0	15.6
6	470	16.0	13.8	4.9	16.9
7	400	13.6	11.5	5.3	23.5
8	205	7.0	5.0	3.5	21.4
9	233	7.9	5.3	4.0	17.7
10					
Group B					
1	317	10.8	7.0	5.8	18.5
2	354	12.0	7.5	7.8	27.5
3	306	10.4	8.0	5.7	31.7
4	396	13.5	11.3	4.8	19.3
5	290	9.9	6.5	4.8	14.1
6	333	11.3	9.6	7.3	49.5
7	335	11.4	9.7	4.6	25.4
8	273	9.3	7.9	4.7	35.5
9	215	7.3	5.3	3.8	24.6
10	220	7.5	5.6	3.4	20.0
Group C					
1	337	11.5	7.0	8.0	30.4
2	387	13.2	9.1	7.5	25.7
3	387	13.2	9.3	7.8	29.5
4	222	7.5	5.4	3.6	20.0
5	350	11.9	8.7	5.0	15.1
6	398	13.5	9.6	5.9	14.8
7	390	13.3	103	7.7	45.6
8	230	7.8	5.2	3.7	14.1
9	320	10.9	8.3	4.1	13.7
10	223	7.6	5.1	3.4	11.8

The mild hypothermia (Group 2) and moderate hypothermia (Group 3) groups showed decreased caspase-9 enzyme activity and increased PCNA activity compared with the normothermia group (Group 1). On the other hand, there was no statistically significant difference between the moderate hypothermia (Group 3) and mild hypothermia (Group 2) groups in terms of caspase-9 activity, whereas moderate hypothermia (Group 3) significantly increased PCNA activity compared with the mild hypothermia group (Group 2) (Figs. 1, 2).

DISCUSSION

In this study, the effects of mild and moderate systemic hypothermia after major hepatectomy on liver damage, oxidative stress, apoptosis, and regenerative capacity were inves-

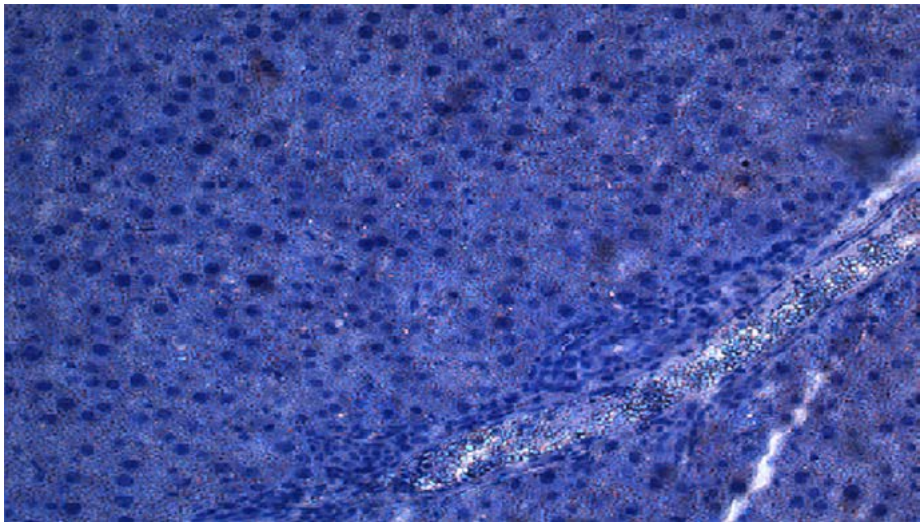


Figure 1. In the picture of the moderate hypothermia group, mild caspase-9 uptake is observed in hepatocytes. (Bar: 2 µm, Anti-caspase-9, immunoperoxidase, X 400)

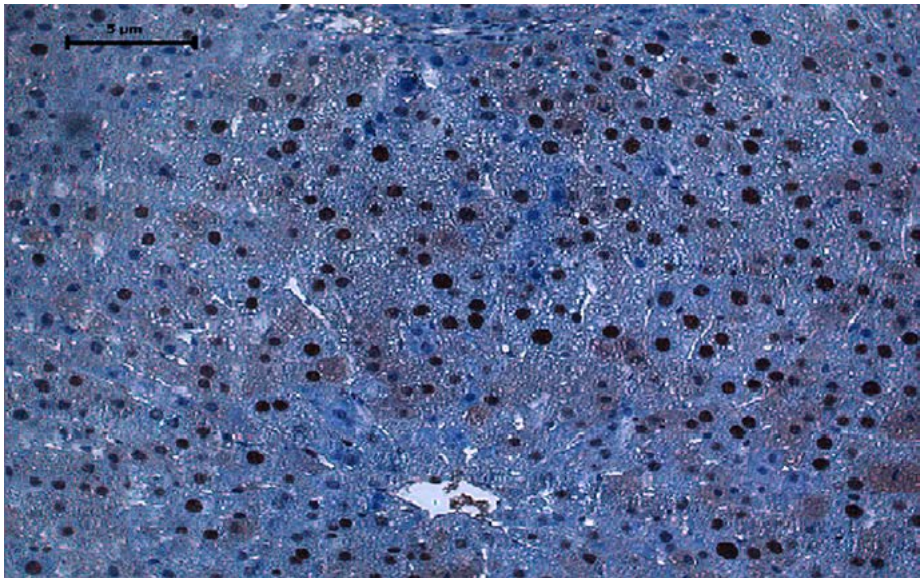


Figure 2. In the picture of the mild hypothermia group, hepatocytes arranged radially around the central vein show moderate to intense uptake (bar: 5µm, PCNA Immunoperoxidase, X 400)

tigated, and it was revealed that the effects of hypothermia differ depending on temperature.

Liver regeneration, which occurs through compensatory hyperplasia and hypertrophy of the residual liver lobes without the involvement of stem cells, depends on neoangiogenesis and tissue proliferation.[12] When a trigger such as hepatectomy or liver damage occurs, almost all (95%) liver cells re-enter the cell cycle (S phase, 12–24 hours), triggering DNA synthesis; DNA synthesis also occurs in nonparenchymal cells. DNA synthesis progresses from the periportal area toward the central vein. Most of the increase in liver mass occurs within the first 3 days after PHx, and parenchymal restoration is completed within 5–7 days.[13] Experimental studies have shown that, in addition to many factors regulating

tissue restoration following hepatocyte necrosis and partial loss caused by severe viral or drug-induced hepatitis, factors such as serum zinc levels also affect liver regeneration.[14]

Mild hypothermia is used clinically because of its various metabolic and hemodynamic effects. It has been used as part of the treatment for increased intracranial pressure and has led to a decrease in mortality in an experimental hemorrhagic shock model in rats.[15,16] The purpose of hypothermia in liver surgery is to increase the duration and tolerance of ischemia. The most important cause of mortality after liver resection is operative blood loss. Clamping techniques such as Pringle and TVE are used to prevent this blood loss during partial resection. The main problem with the use of these techniques is that the tissue is exposed to I/R injury. With SOR formation during reperfusion, OS increases and mediates liver damage.[17]

Black et al.[18] demonstrated that systemic hypothermia induced by whole-body cooling decreases I/R injury by prolonging the ischemic period following liver resection but also produces side effects, including cardiac arrhythmia, severe hypotension, metabolic acidosis, and coagulopathy. Dinant et al.[17] used hypothermic perfusion to reduce I/R injury during TVI and decreased the liver core temperature by 10°C. In the normothermic group, TVI caused cellular vacuolization, sinusoidal congestion, neutrophil infiltration, and mitochondrial dysfunction. In the hypothermic perfusion group, liver function indicators such as AST levels and histological damage scores decreased significantly, and survival increased. Fortner et al.[19] performed major hepatic resection for liver tumors with hypothermic perfusion under TVE, with a mortality rate of 10%, whereas Yamanaka et al.[20] performed segmentectomy in cirrhotic livers with hepatocellular carcinoma under mild hypothermic conditions provided by hemihepatic vascular occlusion and topical cooling. They showed that ischemia duration was prolonged, hepatic function was preserved, and complications were reduced with hemihepatic cooling and occlusion compared with the control group. Vaquero et al.[21] showed that, in APAP-induced liver damage in rats, apoptotic cell death and mortality decreased under mild hypothermic conditions compared with normothermia. Grezzana Filho et al.[22] found that selective hepatic hypothermia prevented I/R injury by inhibiting the release of inflammatory cytokines, protected the microcirculation, prevented hepatocellular necrosis and leukocyte infiltration, and allowed preservation of liver architecture.

The effects of hypothermia on liver regeneration are controversial. According to Munoz, hypothermia negatively affects liver regeneration because of decreased cardiac output, mesenteric and hepatic blood flow, changes in genes involved in regeneration, and possible downregulation of biochemical enzymes involved in the regeneration process.[23] In an experimental study investigating the effect of partial liver resection performed under hypothermic conditions on regeneration, Arab et al.[16] divided mice into 2 groups: mild hypothermia and normothermia. The temperature of the hypothermia group was maintained at 32°C, and partial liver resection (2/3) was performed in both groups for 3 hours under standardized conditions, taking the diurnal rhythm into consideration. Although a strong regenerative response was observed in both groups, no difference was found between the groups in morphological and immunohistochemical evaluations. Hypothermia had no effect on regeneration in mice subjected to partial liver resection.

In our study, although there was an increase in the regenerative response evaluated by PCNA staining, which is a standard method, this effect was not observed in morphological measurements in the hypothermia groups. This may be because the increased mitotic activity at 48 hours had not yet resulted in a significant increase in the remaining liver mass. In addition, unlike the study by Arab et al.,[16] our finding of an increased regenerative response with greater hypothermia depth may be related to the fact that the diurnal rhythm was not taken into account during resections or to the duration of hypothermia. While hypothermia may have a positive effect on regeneration over short periods, it may suppress the response because of possible hemodynamic effects when the duration is prolonged. Studies on hypothermia models of different durations and depths, in which possible effects such as diurnal rhythm are standardized, are needed.

Hypothermia, in addition to slowing metabolism and delaying the damage process caused by oxygen and substrate deficiencies, may itself increase cellular damage. This damage is mediated by changes in ion homeostasis and membrane fluidity. The effect of hypothermia on necrosis and apoptosis, which are forms of cellular damage, is controversial, and studies in the literature have shown effects in both directions. Rauen et al.[24] showed that apoptosis occurs in both normal and hypothermic cell cultures, but most apoptotic indicators increase with the duration of hypothermia and especially during rewarming, which they described as cold-induced. The researchers drew attention to the activity of SOR in the mechanism of damage and demonstrated this activity by showing that apoptosis was almost completely suppressed by the addition of antioxidants to the medium. Fu et al.[25] showed that Fas ligand-mediated apoptosis was suppressed

through partial inhibition of signaling events, including mitochondrial damage, cytochrome c release, subsequent apoptosis, and effector caspase activation, in mouse hepatocytes cultured at 32°C. Molodykh et al.[26] showed that hepatocyte necrosis developed as a result of reorganization of stromal and parenchymal cells in rats subjected to whole-body hypothermia. In our study, biochemical and histological indicators of liver damage did not differ among the groups. Hypothermia did not affect liver injury. In the literature, studies in which damage parameters were evaluated after hypothermic liver surgery have primarily involved I/R models.

The reason why liver injury parameters were not positively or negatively affected by hypothermia in our study may be the absence of an ischemic period. Since continuous hemodynamic monitoring was not performed during our study, it is not possible to draw a clear conclusion on this issue. Previous studies have shown that hypothermia increases oxidative stress in both liver tissue and other organs.[18,24] In our study, hypothermia was also found to increase oxidative stress. There is strong evidence of a correlation between oxidative stress levels and inflammatory cytokines.[27] In our study, systemic mild hypothermia was also shown to increase oxidative stress after hepatectomy while suppressing apoptosis. Moderate hypothermia caused oxidative stress levels to return to normal. Despite the decrease in oxidative stress, no change was observed in the apoptosis-suppressing effect of moderate systemic hypothermia compared with mild hypothermia.

Hypothermia may exert its apoptosis-suppressing and hepatocyte-protective effects by decreasing proinflammatory cytokines.

In our study, mild hypothermia increased oxidative stress, as reported in a previous study.[28] However, unlike the findings of that study, mild hypothermia in our study favorably contributed to regeneration by increasing hyperplasia. In addition, mild hypothermia was shown to suppress apoptosis without affecting liver damage. Experimental studies using appropriate protocols are needed to understand the mechanisms underlying this chain of events. Qin et al.[29] showed that mild hypothermia preconditioning can increase liver cell viability, including cell proliferation, LDH regulation, and glycogen synthesis capacity. In addition, the researchers suggested that mild hypothermia preconditioning maintained its protective effect on liver cells in a cold environment. Ohashi et al.[30] emphasized that hypothermia may be a useful clinical marker for the early detection of liver failure after hepatectomy.

When moderate hypothermia was applied, the level of oxidative stress was similar to that observed under normothermic conditions, in contrast to mild hypothermia. Apoptosis was suppressed to a similar extent as with mild hypothermia. However, the increase in liver regeneration was more pronounced than that observed with both normothermia and mild hypothermia. This may be due to the favorable effect of moderate hypothermia without inducing severe oxidative stress.

CONCLUSION

In this study, hypothermia was shown to favorably affect liver regeneration, with a more pronounced effect under moderate hypothermia. The oxidative stress level, which increased with mild hypothermia, was similar to that observed under normothermia with moderate hypothermia. Histological and biochemical parameters showed that hypothermia did not affect liver damage. Hypothermia suppressed apoptosis regardless of its degree.

Disclosures

Ethics Committee Approval: The study was approved by the Başkent University Ethics Committee (No: 2008/AP-455, Date: 29/04/2008).

Informed Consent: Informed consent was obtained from all participants.

Conflict of Interest Statement: The authors have no conflicts of interest to declare.

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

Use of AI for Writing Assistance: In this study, artificial intelligence tools were used for language editing and to reduce the reference list to 250 words.

Author Contributions: Concept – M.Ö., A.Ö.D.; Design – M.Ö., A.Ö.D.; Supervision – R.H.G., P.E.E.; Funding – A.Ö.D., R.A.; Materials – A.Ö.D., R.A.; Data collection and/or processing – A.Ö.D., S.G.B., S.D.D., N.L.Ü., R.A.; Data analysis and/or interpretation – A.Ö.D., M.Ö., P.E.E.; Literature search – A.Ö.D., M.Ö.; Writing – A.Ö.D., M.Ö.; Critical review – R.H.G., P.E.E.

Acknowledgments: We would like to thank Dr. Şebnem Kodsak for her contributions to this study. Although Dr. Kodsak's name was included in the ethics committee file and the proceeding version of research, she was not included among the authors because we were unable to contact her during the manuscript preparation process.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Michalopoulos GK, Bharat B. Liver regeneration: biological and pathological mechanisms and implications. Nat Rev Gastroenterol Hepatol 2021;18:40–55. [CrossRef]
2. Abu Rmilah A, Zhou W, Nelson E, Lin L, Amiot B, Nyberg SL. Understanding the marvels behind liver regeneration. Wiley Interdiscip Rev Dev Biol 2019;8:e340. [CrossRef]
3. Olthof PB, Reiniers MJ, Dirkes MC, Gulik TMV, Golen RFV. Protective mechanisms of hypothermia in liver surgery and transplantation. Mol Med 2016;21:833–46. [CrossRef]
4. Mito M, Tamaki A, Kon T, Ohira S, Mikami J. Experimental studies on differential hypothermia of the liver. J Surg Res 1965;5:207–12. [CrossRef]
5. Yamaoka Y, Ozawa K, Kumada K, Shimahara Y, Tanaka K, Mori K, et al. Total vascular exclusion for hepatic resection in cirrhotic patients. Application of venovenous bypass. Arch Surg 1992;127:276–80. [CrossRef]
6. Hansen TN, Dawson PE, Brockbank KG. Effects of hypothermia upon endothelial cells: mechanisms and clinical importance. Cryobiology 1994;31:101–6. [CrossRef]
7. Ohkawa H, Ohishi N, Yagi K. Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction. Anal Biochem 1979;95:351–8. [CrossRef]
8. Fang YZ, Yang S, Wu G. Free radicals, antioxidants, and nutrition. Nutrition 2002;18:872–9. [CrossRef]
9. Yoshioka T, Kawada K, Shimada T, Mori M. Lipid peroxidation in maternal and cord blood and protective mechanism against activated-oxygen toxicity in the blood. Am J Obstet Gynecol 1979;135:372–6. [CrossRef]
10. Sedlak J, Lindsay RH. Estimation of total, protein-bound, and nonprotein sulfhydryl groups in tissue with Ellman's reagent. Anal Biochem 1968;25:192–205. [CrossRef]
11. Erpek S, Bilgin MD, Doger FK. Elektromanyetik alanın (50HZ, 6mT) sıçan karaciğer ve böbreğine etkileri. ADU Tıp Fak Derg 2007;8:5–11. [Article in Turkish]
12. Andersen KJ, Knudsen AR, Kannerup AS, Sasanuma H, Nyengaard JR, Hamilton-Dutoit S, et al. The natural history of liver regeneration in rats: description of an animal model for liver regeneration studies. Int J Surg 2013;11:903–8. [CrossRef]
13. Moris D, Vernadakis S, Papalampros A, Vailas M, Dimitrokallis N, Petrou A, et al. Mechanistic insights of rapid liver regeneration after associating liver partition and portal vein ligation for stage hepatectomy. World J Gastroenterol 2016;22:7613–24. [CrossRef]
14. Takahashi Y, Oyama H, Nakamura A, Minegishi Y, Tanaka K. Zinc supports liver regeneration after partial resection. Turk J Surg 2023;39:344–53. [CrossRef]
15. Choi S, Noh J, Hirose R, Ferell L, Bedolli M, Roberts JP, et al. Mild hypothermia provides significant protection against ischemia/reperfusion injury in livers of obese and lean rats. Ann Surg 2005;241:470–6. [CrossRef]
16. Arab JP, Pizarro M, Solis N, Sun H, Thevananther S, Arrese M. Mild hypothermia does not affect liver regeneration after partial hepatectomy in mice. Liver Int 2009;29:344–8. [CrossRef]
17. Dinant S, van Veen SQ, Roseboom HJ, van Vliet AK, van Gulik TM. Liver protection by hypothermic perfusion at different temperatures during total vascular exclusion. Liver Int 2006;26:486–93. [CrossRef]
18. Black PR, van Devanter S, Cohn LH. Effects of hypothermia on systemic and organ system metabolism and function. J Surg Res 1976;20:49–63. [CrossRef]
19. Fortner JG, Shiu MH, Kinne DW, Kim DK, Castro EB, Watson RC, et al. Major hepatic resection using vascular isolation and hypothermic perfusion. Ann Surg 1974;180:644–52. [CrossRef]
20. Yamanaka N, Furukawa K, Tanaka T, Tanaka W, Yamanaka J, Imakita M, et al. Topical cooling-assisted hepatic segmentectomy for cirrhotic liver with hepatocellular carcinoma. J Am Coll Surg 1997;184:290–6.
21. Vaquero J, Bélanger M, James L, Herrero R, Desjardins P, Côté J, et al. Mild hypothermia attenuates liver injury and improves survival in mice with acetaminophen toxicity. Gastroenterology 2007;132:372–83. [CrossRef]
22. Grezzana Filho TJM, Longo L, Santos JLD, Gabiatti G, Boffil C, Santos EBD, et al. Induction of selective liver hypothermia prevents significant ischemia/reperfusion injury in Wistar rats after 24 hours. Acta Cir Bras 2020;35:e202000205. [CrossRef]
23. Munoz SJ. Hypothermia may impair hepatic regeneration in acute liver failure. Gastroenterology 2005;128:1143–4; author reply 1144–5. [CrossRef]
24. Rauen U, Polzar B, Stephan H, Mannherz HG, de Groot H. Cold-induced apoptosis in cultured hepatocytes and liver endothelial cells: mediation by reactive oxygen species. FASEB J 1999;13:155–68. [CrossRef]
25. Fu T, Blei AT, Takamura N, Lin T, Guo D, Li H, et al. Hypothermia inhibits Fas-mediated apoptosis of primary mouse hepatocytes in culture. Cell Transplant 2004;13:667–76. [CrossRef]
26. Molodykh OP, Klinnikova MG, Lushnikova EL, Nepomnyashchikh LM. Tissue and intracellular reorganization of rat liver during total body hypothermia. Bull Exp Biol Med 2005;139:741–6. [CrossRef]
27. Yamanaka N, Okamoto E, Furukawa K, Oriyama T, Fujimoto J, Kanno H, et al et al. Hepatic resection under in situ hemihepatic hypothermic perfusion with hepatoprotective agents. Hepatogastroenterology 1995;42:1–4.
28. Cosgrove BD, Cheng C, Pritchard JR, Stolz DB, Lauffenburger DA, Griffith LG. An inducible autocrine cascade regulates rat hepatocyte proliferation and apoptosis responses to tumor necrosis factor-alpha. Hepatology 2008;48:276–88. [CrossRef]
29. Qin J, Mai Y, Li Y, Jiang Z, Gao Y. Effect of mild hypothermia preconditioning against low temperature (4°C) induced rat liver cell injury *in vitro*. PLoS One 2017;12:e0176652. [CrossRef]
30. Ohashi N, Hori T, Uemoto S, Jermanus S, Chen F, Nakao A, et al. Hypothermia predicts hepatic failure after extensive hepatectomy in mice. World J Hepatol 2013;5:170–81. [CrossRef]

DOI: 10.14744/cm.2026.09226
Compreh Med 2026;18(Suppl. 1):25-31

RESEARCH ARTICLE

Phenotypic Characteristics and Co-sensitization Patterns in Pediatric *Cupressus Arizonica* Sensitization

Hande Yüksel Bulut, Enes Çelik

Department of Pediatric Allergy and Immunology, University of Health Sciences Atatürk Sanatorium Training and Research Hospital, Ankara, Türkiye

ABSTRACT

Objective: Sensitization to *Cupressus arizonica* has increasingly been reported in children and may contribute to allergic diseases. This study aimed to evaluate the demographic and clinical characteristics of pediatric patients with *C. arizonica* sensitization and compare monosensitized and polysensitized phenotypes.

Materials and Methods: This retrospective study included 163 children aged 4–18 years with *C. arizonica* sensitization identified by extract-based skin prick testing (SPT). Patients were divided into two groups, monosensitized and polysensitized, based on their SPT and/or specific IgE results.

Results: Among 1560 children undergoing SPT, 163 (10.4%) showed *C. arizonica* sensitization; 38 (23.3%) were monosensitized, and 125 (76.7%) were polysensitized. Ninety-eight (60.1%) patients were male. The median current age was 136 months (IQR, 100–180). Polysensitized patients had an earlier symptom onset (p=0.03), while age, sex, and symptom patterns were comparable between the groups. Allergic rhinitis was significantly more frequent in polysensitized children (84.8% vs. 52.6%, p=0.01), whereas the presence of asthma, atopic dermatitis, chronic urticaria/angioedema, disease severity, lung function parameters, and environmental factors did not differ significantly. Total IgE levels and eosinophil counts were significantly higher in the polysensitized group (p=0.01 for both). *C. arizonica* SPT wheal and erythema diameters were similar between the groups. Among polysensitized patients, co-sensitization was most common for timothy grass (75.2%), cultivated grasses (65.6%), and cat dander (54.4%).

Conclusion: *C. arizonica* sensitization is increasingly prevalent and frequently accompanied by co-sensitization. Although polysensitized patients showed earlier symptom onset and more frequent allergic rhinitis, the number of sensitizations was not directly associated with disease severity, highlighting the multifactorial nature of allergic diseases.

Keywords: Allergic rhinitis, asthma, atopic dermatitis, children, *Cupressus arizonica*, sensitization

How to cite this article: Yüksel Bulut H, Çelik E. Phenotypic Characteristics and Co-sensitization Patterns in Pediatric *Cupressus Arizonica* Sensitization. Compreh Med 2026;18(Suppl. 1):25-31.

INTRODUCTION

Childhood allergic diseases are strongly influenced by local aeroallergen exposure.[1] Sensitization to tree pollens from the Cupressaceae family has been increasingly reported in pediatric populations, particularly in Mediterranean and temperate regions, owing to their prolonged pollination periods and expanding distribution.[2–4] In Türkiye, tree pollens constitute a substantial component of the pollen burden, with Cupressaceae pollens detectable across extended periods of the year.[5] Among these, *Cupressus arizonica* has emerged as a clinically relevant aeroallergen due to its environmental resilience and widespread use in urban landscaping.[2,3,6] Climate change–related alterations in pollen distribution and seasonality may further contribute to rising sensitization rates and the growing burden of associated allergic diseases.[1–3] Recent evidence suggests that up to 40% of individuals sensitized to Cupressaceae develop clinically relevant allergy, predominantly presenting as rhinitis or rhinoconjunctivitis, while allergic diseases such as asthma and atopic dermatitis also occur.[2,4,6,7]

Address for Correspondence: Hande Yüksel Bulut, Department of Pediatric Allergy and Immunology, University of Health Sciences Atatürk Sanatorium Training and Research Hospital, Ankara, Türkiye
E-mail: handeyuksel_88@hotmail.com **ORCID ID:** 0000-0001-6569-9877

Received date: 11.05.2026
Revised date: 03.07.2026
Accepted date: 16.07.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Previous studies on aeroallergen sensitization patterns have demonstrated that polysensitization is associated with more persistent symptoms, higher total IgE levels, and multiple comorbidities.[8–10] However, it remains unclear whether these associations reflect distinct clinical and laboratory phenotypes in mono- and polysensitized children with *C. arizonica*. Establishing this distinction is clinically important for risk stratification and individualized treatment.

In the Central Anatolia region, where local aerobiology and exposure profiles may differ from those in Mediterranean coastal regions, studies evaluating children with *C. arizonica* sensitization are limited in the literature. Accordingly, this study aimed to evaluate the demographic, clinical, and laboratory characteristics of pediatric patients with extract-defined *C. arizonica* sensitization in the Central Anatolia region of Türkiye, to compare mono- and polysensitized phenotypes, and to characterize their co-sensitization patterns.

MATERIALS and METHODS

Study Population and Design

This retrospective observational study was conducted at the Pediatric Allergy Outpatient Clinic of a tertiary care center. Medical records of 1560 patients aged 4–18 years who were diagnosed with allergic diseases, including asthma, allergic rhinitis, atopic dermatitis, and chronic urticaria, and who underwent skin prick testing between January 2024 and January 2025 were reviewed. Among these, 163 patients with *C. arizonica* sensitization identified by extract-based SPT were included in the study. Patients without a confirmed diagnosis of allergic disease or with incomplete demographic, clinical, or laboratory data were excluded.

Demographic and clinical data, including age, sex, clinical symptoms and severity, age at symptom onset, symptom pattern (seasonal vs. perennial), environmental exposures, and family history of allergic diseases, were extracted from the patients' electronic medical records.

Patients with allergic rhinitis and/or allergic rhinoconjunctivitis were evaluated for nasal and ocular symptoms at presentation according to the Allergic Rhinitis and Its Impact on Asthma (ARIA)-EAACI Guidelines.[11] Asthma was diagnosed according to the 2025 Global Initiative for Asthma (GINA) Strategy Report.[12] The evaluated clinical parameters included daytime and nocturnal symptoms, exercise-induced symptoms, the number of exacerbations in the previous year, and a history of hospitalization due to asthma exacerbations. In patients aged 6 years and older, lung function tests were performed, and forced expiratory volume in one second (FEV_1) and bronchodilator reversibility percentages were recorded. Chronic urticaria was diagnosed according to the International EAACI/GA²LEN/EuroGuiDerm/APAAACI Guideline for the Definition, Classification, Diagnosis, and Management of Urticaria.[13] The severity of atopic dermatitis was assessed using the Patient-Oriented SCORAD (PO-SCORAD) index.[14]

All patients underwent skin prick testing (SPT) with standardized aeroallergen extracts using two panels containing a total of 16 aeroallergens, administered at least two weeks apart. Skin prick tests were performed on the volar surface of the forearm using a standardized 1-mm lancet. Wheal diameters were measured after 20 minutes, and a wheal size ≥3 mm greater than the negative control was considered positive. Skin prick tests were performed using standardized commercial allergen extracts obtained from ApiPrick® (Asacpharma, Spain) and Lofarma® (Milan, Italy). In patients with asthma, SPTs were performed during periods of well-controlled asthma. In patients with atopic dermatitis, SPT was applied to eczema-free skin areas. In the chronic urticaria group, SPTs were conducted when symptoms were controlled and the patients were not taking antihistamine therapy.

The allergen panel included *C. arizonica*, timothy grass, cultivated grasses (*Hordeum, Secale, Triticum*), a tree pollen mixture (*Alnus, Corylus, Betula, Fraxinus*), *Ambrosia artemisiifolia, Parietaria, Plantago lanceolata, Artemisia vulgaris, Chenopodium*, a *Dermatophagoides* mixture, cockroach, cat dander, dog dander, *Alternaria alternata, Aspergillus fumigatus*, and *Cladosporium*.

Peripheral blood eosinophil count and percentage, total IgE, and allergen-specific IgE (sIgE) levels (grass pollen mixture, dust mixture, mold mixture, and cat dander) were measured in all patients. The grass pollen mixture sIgE included timothy grass, perennial rye grass, orchard grass, meadow fescue, and June grass. The dust mixture sIgE contained the allergens *Dermatophagoides pteronyssinus, Dermatophagoides farinae*, cockroach, and house dust-greer. The mold mixture sIgE contained *Alternaria tenuis, Cladosporium herbarum, Aspergillus fumigatus, Candida albicans*, and *Penicillium notatum*. Specific IgE levels were measured using a chemiluminescent immunoassay. The lower detection limit for all sIgE assays was 0.35 kIU/L.

Patients were classified into two groups based on SPT and/or sIgE results: a monosensitized group and a polysensitized group. Monosensitization was defined as sensitization exclusively to *C. arizonica*, demonstrated by a positive SPT to *C. arizonica*, with negative SPT results for all other aeroallergens in the 16-allergen panel and negative sIgE results (<0.35 kIU/L)

for all tested allergen mixtures and cat dander. Polysensitization was defined as sensitization to *C. arizonica* plus at least one additional aeroallergen, as determined by SPT and/or sIgE. Co-sensitization was defined as concomitant sensitization to *C. arizonica* and one or more additional aeroallergens.

Ethical approval was obtained from the hospital's ethics committee on September 24, 2025, with protocol number 2024-BÇEK/351. This study was conducted in accordance with the Declaration of Helsinki.

Statistical Analysis

Statistical analysis was performed using IBM SPSS Statistics (Version 26.0; IBM Corporation, Armonk, New York, USA). Descriptive statistics, including frequency and percentage distributions, were used to analyze the data. The normality of the data distribution was evaluated using the Shapiro–Wilk test. The chi-square test or Fisher's exact test was used to compare categorical variables. The Mann–Whitney U test was used to compare quantitative data between two groups. Statistical significance was accepted as $p<0.05$.

RESULTS

A total of 163 (10.4%) patients with sensitization to *C. arizonica* were identified among 1560 individuals who underwent SPT. Among sensitized patients, 38 (23.3%) were monosensitized and 125 (76.7%) were polysensitized.

Ninety-eight (60.1%) of the patients were male. Sex distribution did not differ between the monosensitized and polysensitized groups (p=0.19). Median current age was also similar (122.5 months [IQR 96.7–165] vs. 138 months [IQR 101.5–180.5], p=0.34). In contrast, polysensitized patients had an earlier age at symptom onset than monosensitized patients (median 89 months [IQR 72–101.5], p=0.03). Among patients sensitized to *C. arizonica*, 54.6% reported seasonal symptoms, whereas 45.4% reported perennial symptoms. Symptom patterns (seasonal or perennial) were comparable between the groups (p=0.46). Detailed demographic and clinical characteristics are summarized in Table 1.

Allergic rhinitis/rhinoconjunctivitis was the most frequent allergic disease and was more common in polysensitized patients (84.8%) than in monosensitized patients (52.6%, p=0.01). In contrast, the frequencies of asthma, atopic dermatitis, and chronic urticaria were similar between the groups (all p>0.05). The number of allergic diseases (single vs. multiple) did not differ significantly between the monosensitized and polysensitized groups (p=0.33). Current allergic rhinitis/rhinoconjunctivitis and asthma symptoms, the number of asthma exacerbations per year, hospitalization due to exacerbations, and lung function parameters (FEV_1 and reversibility percentage) were also comparable between monosensitized and polysensitized patients (all reported p>0.05, Table 1). Among patients with atopic dermatitis, median PO-SCORAD scores were comparable between the monosensitized and polysensitized groups (p=0.81).

Environmental and familial factors, including pet ownership, tobacco smoke exposure, and family history of allergy, did not differ significantly between the groups (all p>0.05).

C. arizonica SPT wheal diameters did not differ significantly between monosensitized and polysensitized children (p=0.66). Total IgE levels were higher in polysensitized patients (median 259.5 IU/mL [IQR 130–585.4]) than in monosensitized patients (median 85 IU/mL [IQR 31–137], p=0.01). Absolute eosinophil counts and percentages (median 360 [IQR 180–480] and 3.5% [IQR 2.1–5.7], respectively) were higher in the polysensitized group (p<0.05 for both). The distribution of SPT and laboratory parameters according to sensitization status is shown in Table 2.

Among polysensitized children, co-sensitization was most frequently observed for timothy grass (75.2%), followed by cultivated grasses (65.6%) and cat dander (54.4%). Sensitization to other allergens—including weed and other tree pollens, house dust mites, and molds—occurred less frequently. The co-sensitization profile of polysensitized patients is presented in Table 3.

Median allergen-specific IgE levels differed markedly between monosensitized and polysensitized patients. In the monosensitized group, median sIgE levels were 0.1 kU/L [0.1–0.2] for grass mixture, dust mixture, and mold mixture and 0.1 kU/L [0.1–0.1] for cat dander. In polysensitized patients, median sIgE levels were 63.7 kU/L [20.9–100] for grass mixture in those with grass/weed/tree pollen sensitization, 4.6 kU/L [1.3–15.6] for dust mixture in those with house dust mite sensitization, 2.7 kU/L [1.1–11.9] for mold mixture in those with mold sensitization, and 10.8 kU/L [0.7–95.1] for cat dander in those with cat sensitization.

DISCUSSION

In this study conducted in the Central Anatolia region, sensitization to *C. arizonica* was identified in 10.4% of pediatric patients presenting to a tertiary allergy clinic with allergic symptoms. Our findings indicate that *C. arizonica* sensitization is relatively common among children in Central Anatolia, and the observed prevalence was higher than the 7.5% reported in a pediatric cohort from the same region approximately a decade earlier.[15] This increase may be related to environmental changes, such as an increase in the number of

Table 1. Demographic and clinical characteristics of the study population

	All patients n=163	Monosensitized n=38	Polysensitized n=125	p
Male sex, n (%)	98 (60.1)	20 (52.6)	78 (62.4)	0.19
Current age, median [IQR], months	136 [100–180]	122.5 [96.7–165]	138 [101.5–180.5]	0.34
Age at symptom onset, median [IQR], months	91 [72–114]	98 [72–123.3]	89 [72–101.5]	0.03
Symptom pattern, n (%)				0.46
Seasonal	89 (54.6)	20 (52.6)	69 (55.2)	
Perennial	74 (45.4)	18 (47.4)	56 (44.8)	
Allergic diseases, n (%)				
AR/RC	126 (77.3)	20 (52.6)	106 (84.8)	0.01
Asthma	66 (40.5)	15 (39.5)	51 (40.8)	0.52
AD	41 (25.1)	12 (31.6)	29 (23.2)	0.20
CU	19 (11.7)	7 (18.4)	12 (9.6)	0.12
Number of Allergic Diseases, n (%)				0.33
Single allergic disease	87 (53.4)	22 (57.9)	65 (52.0)	
Multiple allergic diseases	76 (46.6)	16 (42.1)	60 (48.0)	
AR/RC + Asthma	36 (21.8)	5 (13.2)	31 (24.8)	
AR/RC + AD	17 (10.3)	4 (10.5)	13 (10.4)	
Asthma + AD	5 (3.0)	4 (10.5)	1 (0.8)	
AR/RC + Asthma + AD	9 (5.5)	–	9 (7.2)	
CU + additional allergic diseases	9 (5.5)	3 (7.9)	6 (4.8)	
Current AR/RC symptoms, n (%)	113 (69.3)	20 (52.6)	93 (74.4)	0.28
Rhinorrhea	80 (49.1)	11 (28.9)	69 (55.2)	0.20
Itchy nose	83 (50.9)	16 (42.1)	67 (53.6)	0.19
Nasal congestion	88 (54.0)	16 (42.1)	72 (57.6)	0.32
Sneezing	84 (51.5)	14 (36.8)	70 (56.0)	0.59
Eye symptoms	70 (42.9)	12 (31.6)	58 (46.4)	0.53
Current Asthma symptoms, n (%)				
Daytime symptoms	30 (18.4)	6 (15.8)	24 (19.2)	0.43
Nighttime symptoms	31 (19.0)	6 (15.8)	25 (20.0)	0.38
Exercise-related symptoms	50 (30.6)	10 (26.3)	40 (32.0)	0.27
Asthma exacerbations/year, median [IQR]	2 [1–2]	2 [1–2]	2 [1–2]	0.34
Hospitalization due to asthma exacerbations, n (%)	22 (13.5)	4 (10.5)	18 (14.4)	0.38
FEV_1, median [IQR], %	86 [74.5–95.5]	90 [82–100]	85.5 [74–95]	0.27
Bronchodilator reversibility, median [IQR], %	14 [12–17]	12.5 [7.5–15.5]	14 [12–17]	0.24
PO-SCORAD index, median [IQR]	19.5 [6.5–33]	18 [6.8–33]	19.5 [6.1–33.8]	0.81
Any pet ownership, n (%)	45 (27.6)	11 (28.9)	34 (27.2)	0.49
Tobacco exposure, n (%)	96 (58.9)	20 (52.6)	76 (60.8)	0.24
Family history of allergy, n (%)	66 (40.5)	17 (44.7)	49 (39.2)	0.34

IQR: Interquartile range; AR: Allergic rhinitis; RC: Rhinoconjunctivitis; AD: Atopic dermatitis; CU: Chronic urticaria; FEV_1: Forced expiratory volume in one second

Cupressus trees in urban landscapes and changes in pollen exposure due to climate change, as well as differences in the study population.[2,3,5,16] However, because SPT was performed exclusively in patients presenting with allergic symptoms, the observed sensitization rate should not be interpreted as representative of the prevalence in the general population.

Table 2. Median values of the Cupressus arizonica skin prick test wheal diameter and laboratory parameters

	All patients n=163	Monosensitized n=38	Polysensitized n=125	p
SPT wheal diameter, median [IQR], mm	4 [3–5]	4 [3–5]	4 [3–5]	0.66
Total IgE, median [IQR], IU/ml	226 [94.8–565.4]	85 [31–137]	259.5 [130–585.4]	0.01
Eosinophil count, median, [IQR], cells/mm^3	310 [160–470]	180 [140–420]	360 [180–480]	0.01
Eosinophil percentage, median [IQR]	3.3 [1.9–5.3]	2.3 [1.4–4.3]	3.5 [2.1–5.7]	0.02

IQR: Interquartile range; SPT: Skin prick testing

C. arizonica sensitization most commonly presents clinically as rhinitis or rhinoconjunctivitis.[2,3,17] In this study, allergic rhinitis or rhinoconjunctivitis was the most prevalent clinical manifestation among patients sensitized to *C. arizonica* (77.3%), followed by asthma (40.5%), atopic dermatitis (25.1%), and chronic spontaneous urticaria (11.7%). In a study evaluating pediatric patients predominantly residing in the Central Anatolia region of Türkiye, allergic rhinitis was reported in 89.5%, asthma in 35.8%, and atopic dermatitis in 29.8% of children sensitized to *C. arizonica*.[3] Similarly, another study conducted in the Mediterranean region reported a prevalence of allergic rhinoconjunctivitis of 90.5% in individuals with isolated *Cupressus* allergy.[17] The lower prevalence of allergic rhinitis/rhinoconjunctivitis in our cohort compared with that reported in the literature may reflect the inclusion of a broader clinical spectrum, including atopic dermatitis and chronic spontaneous urticaria in addition to respiratory allergies. These findings indicate that *C. arizonica* sensitization may coexist with various allergic diseases beyond rhinitis/rhinoconjunctivitis; however, no direct association with nonrespiratory allergic diseases can be inferred from the present data.

Cupressaceae sensitization rarely occurs in isolation and is typically part of a broader atopic profile.[3,18] Although monosensitization rates to Cupressaceae pollens have been reported to range from 3% to 14.7%, the higher rate observed in our cohort (23.3%) may reflect differences in study population characteristics and regional variations in aeroallergen exposure.[3,18]

Analysis of the co-sensitization profile revealed that the most frequent sensitizations among polysensitized patients were to timothy grass (75.2%), cultivated grasses (65.6%), and cat dander (54.4%). This pattern suggests that *C. arizonica* sensitization frequently coexists with both seasonal and perennial aeroallergens, including indoor allergens such as cat dander, emphasizing the need for comprehensive allergen evaluation. The Eastern Mediterranean Turkish study likewise reported grass, cat, and weed pollens as the most common additional sensitizations.[3]

Table 3. Allergen co-sensitization profile in polysensitized patients

Co-allergen type	n (%)
Timothy	94 (75.2)
Cultivated grasses (Hordeum, Secale, Triticum)	82 (65.6)
Cat dander	68 (54.4)
Ambrosia artemisiifolia	30 (24.0)
Dust mite mix	25 (20.0)
Plantago lanceolata	20 (16.0)
Artemisia vulgaris	19 (15.2)
Tree pollen mixture	17 (13.6)
Chenopodium	16 (12.8)
Alternaria alternata	15 (12.0)
Dog dander	6 (4.8)
Parietaria	4 (3.2)
Cockroach	4 (3.2)
Aspergillus fumigatus	3 (2.4)
Cladosporium	–

This study further demonstrated that, among pediatric patients sensitized to *C. arizonica*, polysensitization was associated with distinct clinical characteristics compared with monosensitization. Polysensitized children exhibited an earlier onset of allergic symptoms and a higher prevalence of allergic rhinitis/rhinoconjunctivitis, accompanied by elevated total IgE levels and eosinophil counts and percentages, indicative of a more pronounced type 2 inflammatory profile.

Previous studies have reported that polysensitization is associated with older age, higher IgE levels, and more persistent and severe symptoms.[8,18,19] In addition, several studies have demonstrated a positive correlation between the number of sensitizing allergens and the severity of IgE-mediated symptoms.[19] However, our findings partially diverge from the existing literature. In our cohort, neither polysensitization nor the number of sensitizing allergens was significant-

ly associated with the severity of asthma, allergic rhinitis/rhinoconjunctivitis, or atopic dermatitis. This may indicate that disease severity is determined not only by the number of sensitizations but also by multifactorial mechanisms such as genetic predisposition and environmental exposure. Furthermore, the similar distribution of seasonal and perennial symptoms between the monosensitized and polysensitized groups may reflect the prolonged and regionally variable pollination period of *C. arizonica*.[2]

The inclusion of a relatively large cohort of pediatric patients with extract-defined *C. arizonica* sensitization in this study increases the clinical significance of the findings by allowing direct comparison of monosensitized and polysensitized phenotypes. Furthermore, as one of the few studies demonstrating an increasing trend in *C. arizonica* sensitization in the Central Anatolia region, this study contributes to the existing literature. However, several limitations should be considered. The lack of component-resolved diagnostics, particularly Cup a 1-specific IgE, limits the ability to distinguish primary sensitization from cross-reactivity, especially in polysensitized patients. Although all patients had allergy diagnoses confirmed by allergy and immunology specialists, the relationship between *C. arizonica* sensitization and clinically relevant allergic symptoms could not be clearly evaluated. The absence of functional validation methods, such as nasal provocation testing, further restricts interpretation. Furthermore, because the definition of monosensitization was based solely on a predefined aeroallergen panel and the available allergen-specific IgE data, the possibility of sensitization to untested allergens represents an additional limitation of the study. Finally, the retrospective, single-center design in a tertiary care setting may introduce referral bias and limit generalizability.

CONCLUSION

In conclusion, *C. arizonica* sensitization is increasingly prevalent in children, and the high rate of co-sensitization underscores the importance of comprehensive aeroallergen assessment in this population. Monosensitization and polysensitization patterns may exhibit distinct clinical characteristics. Although polysensitized patients tend to show earlier symptom onset and a higher prevalence of allergic rhinitis/rhinoconjunctivitis, the lack of a direct association between the number of sensitizations and disease severity underscores the multifactorial nature of allergic diseases. Prospective, multicenter studies incorporating component-resolved diagnostics and functional validation are needed to better delineate the true clinical relevance of *C. arizonica* sensitization and its cross-reactivity profiles.

Disclosures

Ethics Committee Approval: The study was approved by the The Atatürk Sanatorium Training and Research Hospital Ethics Committee (No: 2024-BÇEK/351, Date: 24/09/2025).

Informed Consent: Since the study is retrospective, informed consent was not obtained.

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

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

Use of AI for Writing Assistance: The authors declared that artificial intelligence was not used in the study.

Author Contributions: Concept – H.YB., E.Ç.; Design – H.YB., E.Ç.; Supervision – H.YB., E.Ç.; Funding – H.YB., E.Ç.; Materials – H.YB., E.Ç.; Data collection and/or processing – H.YB., E.Ç.; Data analysis and/or interpretation – H.YB.; Literature search – H.YB., E.Ç.; Writing – H.YB., E.Ç.; Critical review – H.YB., E.Ç.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Epstein TEG, Rorie AC, Ramon GD, Keswani A, Bernstein J, Codina R, et al. Impact of climate change on aerobiology, rhinitis, and allergen immunotherapy: Work Group Report from the Aerobiology, Rhinitis, Rhinosinusitis & Ocular Allergy, and Immunotherapy, Allergen Standardization & Allergy Diagnostics Committees of the American Academy of Allergy, Asthma & Immunology. J Allergy Clin Immunol 2025;155:1767–82.e2. [CrossRef]
2. Charpin D, Pichot C, Belmonte J, Sutra JP, Zidkova J, Chanez P, et al. Cypress Pollinosis: from Tree to Clinic. Clin Rev Allergy Immunol 2019;56:174–95. [CrossRef]
3. Ünsal H, Şahiner ÜM, Soyer Ö, Şekerel BE. Cupressus arizonica: an emerging aeroallergen for East Mediterranean children. Turk J Med Sci 2023;53:1262–70. [CrossRef]
4. Asero R, Ceriotti V, Bonini M. Cypress pollen allergy in Milan: the story of an ongoing growth. Eur Ann Allergy Clin Immunol 2021;53:209–13. [CrossRef]
5. Biçakçi A, Tosunoğlu A. Allergenic pollens in Turkey. Asthma Allergy Immunol 2019;17:7–24. [CrossRef]
6. Chico-Fernández J, Feliu Vila A, Rodríguez-Jiménez B, Valbuena Garrido T, Ayuga-Téllez E. Influence of atmospheric pollutants on allergic sensitization to cupressaceae, Olea, and Platanus pollen in the community of Madrid (2017–2021). Life (Basel) 2025;15:1774. [CrossRef]
7. Charpin D, Calleja M, Pichot C, Penel V, Hugues B, Poncet P. Cypress pollen allergy. Rev Mal Respir 2013;30:868–78. [CrossRef]
8. Bakır DB, Yağmur H, Kabadayı G, Boyacıoğlu ÖK, Atay Ö, Asilsoy S, et al. Demographics and clinical features of pediatric patients with allergic rhinitis: a single-center study from Western Turkey. BMC Pediatr 2025;25:170. [CrossRef]
9. Ciprandi G, Cirillo I. Monosensitization and polysensitization in allergic rhinitis. Eur J Intern Med 2011;22–9. [CrossRef]
10. Raciborski F, Bousquet J, Bousqet J, Namysłowski A, Krzych-Fałta E, Tomaszewska A, et al. Dissociating polysensitization and multimorbidity in children and adults from a Polish general population cohort. Clin Transl Allergy 2019;9:4. [CrossRef]

11. Bousquet J, Sousa-Pinto B, Vieira RJ, Schünemann HJ, Zuberbier T, Bognanni A, et al. Methodology for the development of the Allergic Rhinitis and Its Impact on Asthma (ARIA)-EAACI 2024–2025 Guidelines: from evidence-to-decision frameworks to digitalised shared decision-making algorithms. Allergy 2026;81:427–53. [CrossRef]
12. Global Initiative for Asthma (GINA). 2025 GINA strategy report. Global Initiative for Asthma; 2025. Available at: https://ginasthma.org/2025-gina-strategy-report/. Accessed September 22, 2026.
13. Zuberbier T, Abdul Latiff AH, Abuzakouk M, Aquilina S, Asero R, Baker D, et al. The international EAACI/GA²LEN/EuroGuiDerm/APAAACI guideline for the definition, classification, diagnosis, and management of urticaria. Allergy 2022;77:734–66. [CrossRef]
14. Stalder JF, Barbarot S, Wollenberg A, Holm EA, De Raeve L, Seidenari S, et al. Patient-Oriented SCORAD (PO-SCORAD): a new self-assessment scale in atopic dermatitis validated in Europe. Allergy 2011;66:1114–21. [CrossRef]
15. Cavkaytar O, Buyuktiryaki B, Sag E, Soyer O, Sekerel BE. What we miss if standard panel is used for skin prick testing? Asian Pac J Allergy Immunol 2015;33:211–21. [CrossRef]
16. Rottem M, Noujedat M, Awni Y. Sensitizations to aeroallergens in Israel: Prevalences and profiles. Allergol Immunopathol (Madr) 2025;53:101–11. [CrossRef]
17. Alonso-Díaz de Durana MD, Pérez-Fernández E, Villalba M, Martín-Pedraza L, Fernández-Rivas M. Allergy to cypress and olive pollen: clinical phenotypes and allergen recognition. J Investig Allergol Clin Immunol 2024 3;34:313–22. [CrossRef]
18. Italian Association of Aerobiology. An epidemiological survey of Cupressaceae pollenosis in Italy. J Investig Allergol Clin Immunol 2002;12:287–92.
19. Ha EK, Baek JH, Lee SY, Park YM, Kim WK, Sheen YH, et al. Association of polysensitization, allergic multimorbidity, and allergy severity: a cross-sectional study of school children. Int Arch Allergy Immunol 2016;171:251–60. [CrossRef]

DOI: 10.14744/cm.2026.76094
Compreh Med 2026;18(Suppl. 1):32-37

RESEARCH ARTICLE

Evaluation of Physicians' Knowledge and Proficiency in The Management of Anaphylaxis and The Use of Adrenaline Auto-Injectors

Mehmet Özkaya[1], Hatice Çelik Tuğlu[2]

[1]Department of Pediatric Immunology and Allergy Diseases, Şırnak State Hospital, Şırnak, Türkiye
[2]Department of Immunology and Allergy Diseases, Şırnak State Hospital, Şırnak, Türkiye

ABSTRACT

Objective: This study was conducted to evaluate the knowledge levels of specialist and general practitioners working at Şırnak State Hospital regarding anaphylaxis treatment and their ability to correctly use an adrenaline auto-injector.

Materials and Methods: A questionnaire measuring knowledge levels regarding anaphylaxis treatment was administered to the physicians. The participating physicians were asked to demonstrate the use of the adrenaline auto-injector. Data obtained from the evaluation results were recorded. The use of the adrenaline auto-injector was then explained to the physicians again.

Results: A total of 108 physicians participated in the study, 60 of whom were specialists (55.6%) and 48 were general practitioners (44.4%). Of the participating physicians, 87 (80.6%) correctly knew that adrenaline was the first drug to be administered in anaphylaxis. Of the 108 physicians, 94 (87%) correctly knew the dose of adrenaline. Of the 108 physicians surveyed, 101 (93.5%) correctly knew the route of adrenaline administration, and 92 (85.2%) knew where it should be administered. The steps involved in administering auto-injectors were also evaluated, and incorrect administration steps were identified. The physicians' adrenaline auto-injector administration practices were also assessed after the training.

Conclusion: In conclusion, improvements in the administration of adrenaline auto-injectors by physicians, the development of clinical skills, improvement in physicians' knowledge regarding anaphylaxis and adrenaline administration, and increased training in this area are necessary. It is crucial that physicians have the confidence and competence to administer adrenaline auto-injectors safely.

Keywords: Adrenaline, allergy, anaphylaxis, auto-injector, treatment

How to cite this article: Özkaya M, Çelik Tuğlu H. Evaluation of Physicians' Knowledge and Proficiency in The Management of Anaphylaxis and The Use of Adrenaline Auto-Injectors. Compreh Med 2026;18(Suppl. 1):32-37.

INTRODUCTION

Anaphylaxis is a life-threatening hypersensitivity reaction that begins acutely and involves multiple organ systems.[1] The signs and symptoms of anaphylaxis are varied and can affect the skin, respiratory and gastrointestinal systems, and cardiovascular system.[2] The recommended treatment for anaphylaxis is usually intramuscular administration of adrenaline in the community.[3] Since it is impossible to predict when anaphylaxis will occur in individuals who have experienced it, an adrenaline auto-injector (AAI) is prescribed to the patient. Patients, their parents, and even their teachers should be adequately informed about how to recognize anaphylaxis and how to administer the AAI.[4] Recognizing anaphylaxis and administering adrenaline early are crucial. Delays in the treatment of anaphylaxis can occur, especially in anaphylaxis cases.[5] As a result of delays in administering adrenaline to the patient, the risk of mortality increases.[6] Due to insufficient recognition of the symptoms and signs of anaphylaxis, anxiety and fear experienced during administration, and unfamiliarity with auto-injectors, the prompt use of AAIs may be limited.[7]

Address for Correspondence: Mehmet Özkaya, Department of Pediatric Immunology and Allergy Diseases, Şırnak State Hospital, Şırnak, Türkiye
E-mail: dr.mehmetozk@gmail.com **ORCID ID:** 0000-0002-8101-9058

Received date: 13.05.2026
Revised date: 12.07.2026
Accepted date: 22.07.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Although previous studies, including several reports from Türkiye, have evaluated physicians' knowledge of anaphylaxis, data assessing both theoretical knowledge and practical adrenaline auto-injector (AAI) administration skills, together with the impact of structured hands-on training, remain limited. Therefore, this study aimed to identify gaps in physicians' knowledge and practical skills regarding anaphylaxis management and AAI use while also evaluating the effectiveness of immediate practical training.

MATERIALS and METHODS

The study was conducted at the Pediatric Allergy and Immunology Outpatient Clinic of Şırnak State Hospital. Ethics committee approval was obtained from Şırnak University on November 10, 2025, with approval number E-74546226-622.01-150011. The study was conducted in accordance with the principles of the Declaration of Helsinki. Specialist physicians and general practitioners who agreed to participate were consecutively enrolled in the study. No additional exclusion criteria were applied. To evaluate physicians' knowledge and practical skills regarding anaphylaxis and adrenaline auto-injector (AAI) use, participants completed a structured questionnaire and were then asked to demonstrate the use of an AAI using a trainer device. A standardized checklist was developed to assess each step of AAI administration (Table 1). Each participant's performance was evaluated by a pediatric allergy and immunology specialist. Following the baseline assessment, participants received standardized face-to-face training, including recognition of anaphylaxis, indications for adrenaline administration, correct AAI administration technique, and hands-on demonstration using the trainer device. Participants were then immediately re-evaluated using the same standardized checklist. Errors during AAI administration were recorded before and after the educational intervention. The questionnaire was developed based on previously published studies and current international anaphylaxis guidelines.[3] Before implementation, the content was reviewed by pediatric allergy and immunology specialists to ensure content validity.

Statistical Analysis

Statistical analysis of the data was performed using IBM SPSS Statistics, version 22.0 (IBM Corp., Armonk, NY, USA). In this study, categorical data were expressed as numbers and percentages. Statistical methods were not applied to groups with univariate variables. The relationship between categorical variables was evaluated using the χ^2 test for pairwise group comparisons. For all results, $p<0.05$ was considered statistically significant.

Ethics Committee Approval

This study was approved by the Şırnak University Ethics Committee with decision number E-74546226-622.01-150011 dated 10.11.2025.

RESULTS

A total of 108 physicians (100%) were included in the study. Of the participants, 60 (55.6%) were specialist physicians and 48 (44.4%) were general practitioners. Overall, 70.4% (n=76) of the physicians correctly defined anaphylaxis, whereas 29.6% (n=32) defined it incorrectly. In addition, 80.6% (n=87) of the physicians correctly identified adrenaline as the treatment for anaphylaxis, whereas 19.4% (n=21) answered incorrectly. When asked about the first drug to be administered in anaphylaxis, 80.6% (n=87) of the physicians answered adrenaline, 14.8% (n=16) answered corticosteroids, and 4.6% (n=5) answered antihistamines (Table 2).

Regarding the adrenaline dose, 87% (n=94) of the participants gave the correct answer, whereas 13% (n=14) gave an incorrect answer. Regarding the muscle into which adrenaline should be administered, 85.2% (n=92) responded vastus lateralis, 13.9% (n=15) deltoid, and 0.9% (n=1) gluteus maximus. Regarding the route of adrenaline administration, 93.5% (n=101) reported intramuscular administration and 6.5% (n=7) reported intravenous administration (Table 2).

A total of 24.1% (n=26) of the participants stated that they had previously received training on anaphylaxis, whereas 75.9% (n=82) stated that they had not. The percentage of those who reported having previously administered an AAI was 5.6% (n=6), whereas 94.4% (n=102) stated that they had not. In response to the question, "Which auto-injector is prescribed for which weight?", 38% (n=41) gave the correct answer, whereas 62% (n=67) gave an incorrect answer. The percentage of those who knew how to correctly use the auto-injector was 39.8% (n=43), whereas 60.2% (n=65) did not. The percentage of physicians who were confident in managing anaphylaxis was 30.6% (n=33), whereas 69.4% (n=75) were not confident (Table 2).

Table 1. Steps for administering adrenaline auto-injector

Steps
1. The adrenaline auto-injector is grasped by the wide white body.
2. First, the cap numbered 1 at the bottom is removed.
3. Then, the locking cap numbered 2 at the top is removed.
4. The auto-injector is applied to the vastus lateralis
5. Wait 10 seconds after application.
6. A light massage is applied after the procedure.

Table 2. Characteristics of physicians

Variable	Total: 108 (100%) n (%)	Variable	Total: 108 (100%) n (%)
Doctor		What is the route of adrenaline administration?	
Specialist Physician	60 (55.6)	Intravenous	7 (6.5)
General Practitioner	48 (44.4)	Intramuscular	101 (93.5)
Defining anaphylaxis		Has received prior training in anaphylaxis?	
Correctly	76 (70.4)	Yes	26 (24.1)
Incorrectly	32 (29.6)	No	82 (75.9)
Intended use of adrenaline in clinical practice		Has he/she administered adrenaline auto-injector (AAI) before?	
Yes	87 (80.6)	Yes	6 (5.6)
No	21 (19.4)	No	102 (94.4)
Knowledge of the first-line treatment for anaphylaxis	16 (14.8)	Which adrenaline auto-injector (AAI) is prescribed for what weight?	
Corticosteroid	5 (4.6)	Knows	41 (38)
Antihistamine	87 (80.6)	Doesn't know	67 (62)
Adrenaline		Does he/she know the correct use of adrenaline auto-injector (AAI) ?	
Adrenaline dose		Yes	43 (39.8)
Correctly	94 (87)	No	65 (60.2)
Incorrectly	14 (13)	Is he/she confident in managing anaphylaxis?	
Which muscle is adrenaline administered to?		Yes	33 (30.6)
Deltoid	15 (13.9)	No	75 (69.4)
Vastus lateralis	92 (85.2)		
Gluteus maximus	1 (0.9)		

Table 3. Mistakes made during adrenaline auto-injector administration steps

Steps	Step taken incorrectly before training	Steps taken incorrectly after training
1. Grasp the adrenaline auto-injector by its wide white body.	0	0
2. First, remove the cap number 1 at the bottom.	10	0
3. Then remove the locking cap number 2 at the top.	18	3
4. Apply the auto-injector to the vastus lateralis.	16	0
5. Wait 10 seconds after application.	30	4
6. Apply a light massage after the procedure.	38	9

When comparing the number of incorrect steps in AAI administration before and after training, the number of errors decreased from 10 to 0 for the step of "removing cap number 1 first," from 18 to 3 for the step of "removing the upper locking cap number 2," and from 16 to 0 for the step of "applying the auto-injector to the vastus lateralis muscle." Errors in the step of "waiting 10 seconds after application" decreased from 30 to 4, and errors in the step of "applying a light massage after the procedure" decreased from 38 to 9. No errors were detected before or after training in the step of "grasping the auto-injector by the wide white body" (0–0) (Table 3).

When comparing specialist physicians and general practitioners, the rate of correctly identifying the definition of anaphylaxis was 81.7% (n=49) among specialists and 56.3% (n=27) among general practitioners, and the difference was statistically significant (p=0.004). The rate of correctly identifying the drug to be administered in anaphylaxis was 95.0% (n=57) among specialists and 62.5% (n=30) among general

Table 4. Comparison of specialist physicians and general practitioners

	Specialist physician n (%)	General practitioner n (%)	p
Accurate definition of anaphylaxis	49 (81.7)	27 (56.3)	0.004
Correct identification of first-line drug to be administered in anaphylaxis	57 (95.0)	30 (62.5)	<0.001
Knowledge of correct adrenaline dose	55 (91.7)	39 (81.3)	0.109
Knowledge of correct injection site	56 (93.3)	37 (77.1)	0.015
Knowledge of correct administration route	59 (98.3)	42 (87.5)	0.029
Previous anaphylaxis training	18 (30.0)	8 (16.7)	0.107
Prior history of AAI administration	6 (10.0)	0 (0.0)	0.024
Knowledge of weight-based AAI prescription	34 (56.7)	7 (14.6)	<0.001
Knowledge of correct AAI use	32 (53.3)	11 (22.9)	0.001
Confidence in anaphylaxis management	21 (35.0)	12 (25.0)	0.262

AAI: Adrenaline auto-injector

practitioners (p<0.001). The rates of correctly identifying the adrenaline dose were 91.7% (n=55) among specialists and 81.3% (n=39) among general practitioners, and the difference was not statistically significant (p=0.109). The rate of correctly identifying the muscle into which adrenaline should be administered was 93.3% (n=56) among specialists and 77.1% (n=37) among general practitioners (p=0.015). The rate of correctly identifying the route of administration was 98.3% (n=59) among specialists and 87.5% (n=42) among general practitioners (p=0.029). The rates of prior anaphylaxis training were 30.0% (n=18) among specialists and 16.7% (n=8) among general practitioners, and the difference was not statistically significant (p=0.107). Prior AAI administration was reported by 10.0% (n=6) of specialists and 0% (n=0) of general practitioners (p=0.024). The rate of correctly knowing how to prescribe an AAI was 56.7% (n=34) among specialists and 14.6% (n=7) among general practitioners (p<0.001), whereas the rate of correctly knowing how to use an AAI was 53.3% (n=32) among specialists and 22.9% (n=11) among general practitioners (p=0.001). Self-confidence rates in anaphylaxis management were 35.0% (n=21) among specialists and 25.0% (n=12) among general practitioners, and the difference was not statistically significant (p=0.262) (Table 4).

DISCUSSION

Our study aimed to evaluate physicians' knowledge of emergency intervention for anaphylaxis, their confidence in using adrenaline auto-injectors (AAIs), their competence in performing the correct steps required for AAI administration, and their administration techniques. In our study, the knowledge of general practitioners and specialist physicians at Şırnak State Hospital regarding the use of adrenaline as the first-line drug in anaphylaxis, its route of administration, dosage, site of administration, and the steps involved in using an auto-injector was evaluated. Some deficiencies were identified among physicians regarding their knowledge and skills in AAI administration.

In our study, 70.4% of physicians correctly defined anaphylaxis, 80.6% correctly identified adrenaline as the first-line treatment, and 87.0% correctly identified the appropriate adrenaline dose. However, only 38.0% correctly identified the appropriate weight-based prescription of AAIs. In contrast, a large-scale study by Cambaz Kurt and Kutlu[8] reported that although more than 90% of pediatricians in Türkiye recognized the signs of anaphylactic reactions, only 53.1% knew the correct auto-injector dose. Another study conducted among physicians showed that 93% of physicians chose epinephrine as the first-line treatment for anaphylaxis, but only 49% knew the correct auto-injector dose.[9] These data are consistent with our study findings: physicians generally recognize anaphylaxis and the need for adrenaline treatment, but they lack knowledge about auto-injector dosing and administration. Administering the correct treatment at the correct dose without delay is vital in anaphylaxis. It is known that delayed or incorrect administration of adrenaline increases the risk of a fatal outcome.[10] Song et al.[11] also reported that rapid intramuscular administration of adrenaline is life-saving and that auto-injectors are insufficiently used in practice. In our study, 93.5% of physicians correctly identified the intramuscular route of adrenaline administration.

Studies evaluating physicians' knowledge and practical skills regarding anaphylaxis and AAI use remain limited. Recent studies from Türkiye have also demonstrated im-

portant deficiencies in physicians' knowledge and practical competence. Deligöz et al.[10] reported that although most resident physicians correctly identified adrenaline as the first-line treatment, knowledge regarding anaphylaxis diagnosis, adrenaline dose, route of administration, and patient follow-up remained suboptimal. Emeksiz et al.[12] demonstrated significant improvements in both theoretical knowledge and practical AAI administration skills among pediatric residents after structured training. Similarly, Akova et al.[13] reported persistent deficiencies in residents' knowledge of anaphylaxis despite generally adequate awareness of first-line adrenaline treatment. In our study, only 39.8% of physicians correctly demonstrated AAI use before training, whereas almost all participants corrected their errors after training. One possible reason for the low rate of correct AAI use in our study may be that general practitioners had not yet received specialty training. General practitioners (22.9%) demonstrated significantly lower knowledge of correct AAI use than specialists (53.3%). This finding suggests that limited experience and lack of specialty training may negatively affect preparedness for anaphylaxis management. Arga et al.[14] also reported that many general practitioners were unable to use an AAI correctly and emphasized that current educational methods were insufficient. The fact that 94.4% of the participants in our study had never previously administered an AAI further supports this finding. The observed reduction in administration errors after training suggests that structured practical education may improve physicians' practical AAI administration skills. Similarly, Arga et al.[14] suggested that combining practical sessions with theoretical instruction significantly improves correct AAI use.

In this study, the steps in which AAI use was performed incorrectly were evaluated. The most frequent error was failure to massage the application site after the procedure. In addition, simple but important deficiencies were identified, such as removing the upper and lower caps of the auto-injector in the incorrect order or, in some cases, removing only the lower cap. After the physicians' mistakes were explained using the demo auto-injector, almost all participants were able to perform the procedure correctly and completely after training. Similarly, in a study conducted among pediatric residents when the older AAI model was in use, improvements were observed after training.[15] It should not be forgotten that AAI use may appear easy, but even simple procedures can become complicated in emergency situations. It is particularly important for physicians, patients, and parents caring for children to know how to use an auto-injector correctly. For this reason, training should be repeated regularly. It is important for physicians to provide this training to patients and their families from the outset and reinforce it at every visit.

A major strength of this study is the simultaneous evaluation of both physicians' theoretical knowledge and practical AAI administration skills. In addition, the immediate assessment before and after standardized hands-on training allowed objective evaluation of the effectiveness of the educational intervention. Furthermore, the inclusion of both specialist physicians and general practitioners enabled comparisons across different levels of clinical experience.

This study has several limitations. First, this was a single-center study using a convenience sample of physicians from a secondary care hospital, which may limit the generalizability of the findings. Second, participants were evaluated immediately after training; therefore, long-term retention of knowledge and practical skills could not be assessed. Previous studies have shown that retention of AAI skills declines over time. In one study, the time elapsed since the last AAI training was the most important factor affecting correct AAI use, and training repeated every six months was associated with a 96% probability of correct AAI use.[16] Similarly, a study conducted among intern physicians demonstrated a decline in AAI use skills six months after training.[17] Finally, participation was voluntary, introducing the possibility of selection bias. These findings emphasize the importance of regular reinforcement of practical AAI training for physicians and other healthcare professionals.

In light of these data, the importance of early intervention and correct auto-injector administration in the management of anaphylaxis is once again emphasized. The need for training healthcare professionals in auto-injector use has also been consistently emphasized in the literature.[16]

CONCLUSION

This study identified important deficiencies in physicians' knowledge and practical skills regarding anaphylaxis management and the use of adrenaline auto-injectors (AAIs). Although fundamental knowledge, including the correct adrenaline dose and route of administration, was generally adequate, correct step-by-step AAI administration remained insufficient. General practitioners demonstrated significantly lower levels of knowledge and practical skills than specialists. The observed improvement after standardized hands-on training suggests that practical educational interventions may improve physicians' preparedness and practical skills in AAI administration. Given the life-threatening nature of anaphylaxis, regular, structured, hands-on AAI training should be incorporated into continuing medical education to improve physicians' preparedness, confidence, and ultimately, patient safety.

Disclosures

Ethics Committee Approval: The study was approved by the Şırnak University Ethics Committee (No: E-74546226-622.01-150011, Date: 10/11/2025).

Informed Consent: Written informed consents were obtained from patients who participated in this study.

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

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

Use of AI for Writing Assistance: The authors declared that artificial intelligence was not used in the study.

Author Contributions: Concept – M.Ö., H.Ç.T.; Design – M.Ö., H.Ç.T.; Supervision – M.Ö., H.Ç.T.; Funding – M.Ö., H.Ç.T.; Materials – M.Ö., H.Ç.T.; Data collection and/or processing – M.Ö., H.Ç.T.; Data analysis and/or interpretation – M.Ö., H.Ç.T.; Literature search – M.Ö., H.Ç.T.; Writing – M.Ö.; Critical review – M.Ö., H.Ç.T.

Acknowledgments: We would like to thank and express our gratitude to the medical professionals for their valuable contributions to this study.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Panesar SS, Javad S, de Silva D, Nwaru BI, Hickstein L, Muraro A, et al. The epidemiology of anaphylaxis in Europe: a systematic review. Allergy 2013;68:1353–61.[CrossRef]
2. Soar J, Pumphrey R, Cant A, Clarke S, Corbett A, Dawson P, et al. Emergency treatment of anaphylactic reactions--guidelines for healthcare providers. Resuscitation 2008;77:157–69.[CrossRef]
3. Muraro A, Worm M, Alviani C, Cardona V, DunnGalvin A, Garvey LH, et al. EAACI guidelines: Anaphylaxis (2021 update). Allergy 2022;77:357–77. [CrossRef]
4. Warren CM, Zaslavsky JM, Kan K, Spergel JM, Gupta RS. Epinephrine auto-injector carriage and use practices among US children, adolescents, and adults. Ann Allergy Asthma Immunol 2018;121:479–89.e2.[CrossRef]
5. Grabenhenrich LB, Dölle S, Moneret-Vautrin A, Köhli A, Lange L, Spindler T, et al. Anaphylaxis in children and adolescents: The European Anaphylaxis Registry. J Allergy Clin Immunol 2016;137:1128–37.e1.[CrossRef]
6. Ellis AK, Day JH. Incidence and characteristics of biphasic anaphylaxis: a prospective evaluation of 103 patients. Ann Allergy Asthma Immunol 2007;98:64–9.[CrossRef]
7. Gallagher M, Worth A, Cunningham-Burley S, Sheikh A. Epinephrine auto-injector use in adolescents at risk of anaphylaxis: a qualitative study in Scotland, UK. Clin Exp Allergy 2011;41:869–77.[CrossRef]
8. Cambaz Kurt N, Kutlu NO. Evaluation of pediatricians' awareness about anaphylaxis. Eur Rev Med Pharmacol Sci 2023;27(Suppl 5):53–61.
9. Sipahi Cimen S, Sayili SB. Level of knowledge among healthcare professionals regarding anaphylaxis. Asia Pac Allergy 2022;12:e41.[CrossRef]
10. Deligöz O, Altaş U. Approaches of surgical and internal branch resident physicians to anaphylaxis. J Cukurova Anesth Surg 2023;6:116–23. [CrossRef]
11. Song TT, Worm M, Lieberman P. Anaphylaxis treatment: current barriers to adrenaline auto-injector use. Allergy. 2014;69(8):983–91.[CrossRef]
12. Emeksiz Z, Gungor A, Demirel A, Dibek Mısırlıoğlu E. Knowledge levels of pediatric assistants on anaphylaxis management and adrenaline autoinjector application skills: pretraining and posttraining evaluation. J Contemp Med 2022;12:789–93.[CrossRef]
13. Akova S, Altaş U, Özkars MY, Aşçıoğlu ME. Evaluation of the anaphylaxis-related knowledge of the resident doctors of the department of pediatrics. Jour Umraniye Pediatr 2023;3:39–44.
14. Arga M, Bakirtas A, Catal F, Derinoz O, Harmanci K, Razi CH, et al. Training of trainers on epinephrine autoinjector use. Pediatr Allergy Immunol 2011;22:590–3.[CrossRef]
15. Grouhi M, Alshehri M, Hummel D, Roifman CM. Anaphylaxis and epinephrine auto-injector training: who will teach the teachers? J Allergy Clin Immunol 1999;104:190–3.[CrossRef]
16. Sirin Kose S, Asilsoy S, Tezcan D, Al S, Atay O, Kangalli O, et al. Is there an optimal training ınterval to ımprove the correct use of adrenaline auto-ınjectors? Int Arch Allergy Immunol 2020;181:136–40.[CrossRef]
17. Topal E, Bakirtas A, Yilmaz O, Karagol IH, Arga M, Demirsoy MS, et al. When should we perform a repeat training on adrenaline auto-injector use for physician trainees? Allergol Immunopathol (Madr) 2014;42:472–5.[CrossRef]

DOI: 10.14744/cm.2026.50470
Compreh Med 2026;18(Suppl. 1):38-46

RESEARCH ARTICLE

Intraoperative Blood Loss Is a Surrogate of Surgical Complexity Rather Than an Independent Predictor of Pedicle Screw Loosening: A Matched Cohort Analysis

Evren Sönmez, Lokman Ayhan, Harun Duran

Department of Neurosurgery, University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye

ABSTRACT

Objective: Pedicle screw loosening is a significant mechanical complication after lumbar instrumented fusion. Intraoperative estimated blood loss (EBL) has been associated with loosening; however, whether this reflects an independent biological effect or surgical complexity remains unclear.

Materials and Methods: This retrospective matched cohort study included 160 patients undergoing posterior lumbar instrumented fusion. Eighty patients with screw loosening were matched 1:1 with 80 controls based on age and sex. Demographic, surgical, radiographic, and paraspinal muscle parameters were analyzed. EBL was evaluated using univariate and multivariate logistic regression models adjusted for surgical complexity (operative time, instrumented levels) and biomechanical–biological factors (PI–LL mismatch, multifidus fat infiltration, DEXA T-score). Composite indices were constructed and compared using ROC analysis.

Results: EBL was significantly higher in the loosening group (p=0.001) and was associated with loosening in univariate analysis but not after multivariable conditional logistic regression (p=0.640). Independent factors associated with loosening included lumbar lordosis (p=0.031), multifidus fat infiltration (p=0.009), and DEXA T-score (p=0.001). The surgical complexity index showed moderate discrimination (AUC=0.80), whereas the biomechanical–biological index demonstrated higher discrimination (AUC=0.90).

Conclusion: EBL is associated with pedicle screw loosening but is not an independent predictor, reflecting surgical complexity. Loosening is predominantly driven by biomechanical and biological factors, highlighting the importance of preoperative optimization of alignment and tissue quality.

Keywords: Intraoperative blood loss, lumbar fusion, paraspinal muscle, pedicle screw loosening, spinopelvic alignment

How to cite this article: Sönmez E, Ayhan L, Duran H. Intraoperative Blood Loss Is a Surrogate of Surgical Complexity Rather Than an Independent Predictor of Pedicle Screw Loosening: A Matched Cohort Analysis. Compreh Med 2026;18(Suppl. 1):38-46.

INTRODUCTION

Pedicle screw loosening is a well-recognized complication after lumbar instrumented fusion, with incidence varying by diagnostic criteria and follow-up.[1,2] Although often radiographically silent, it may progress to instability, pseudarthrosis, and revision surgery.[2,3] Identifying predisposing factors remains a central focus in spine research.[3]

Risk factors are broadly categorized as patient- and surgery-related.[4] Patient factors include reduced bone mineral density, sagittal malalignment, and paraspinal muscle degeneration, all influencing the biomechanical environment at the bone–implant interface.[5–7] Surgical factors—such as instrumentation extent, operative duration, and lumbosacral involvement—increase mechanical load and procedural complexity.[2,8] Although individually associated with loosening, their relative contributions and interactions remain unclear.[4,8]

Among perioperative variables, intraoperative estimated blood loss (EBL) is frequently higher in patients who develop loosening.[2,8] This has led to the hypothesis that blood loss may impair local perfusion and osseointegration, contributing to implant failure.[9] However, EBL is not

Address for Correspondence: Evren Sönmez, Department of Neurosurgery, University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
E-mail: evren-sonmez@hotmail.com **ORCID ID:** 0000-0002-0457-201X

Received date: 19.04.2026
Revised date: 12.07.2026
Accepted date: 05.08.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

an independent variable but a consequence of procedural complexity.[8] Larger constructs, longer operative times, and more extensive corrections increase blood loss and are also linked to loosening,[2,8] suggesting confounding rather than causation.[4,8]

Few studies have assessed whether EBL retains independent predictive value after adjustment for surgical complexity and biomechanical vulnerability.[8] Moreover, prior work has focused on individual factors without addressing the hierarchical relationship between surgical burden and the biological substrate. Mechanistically, loosening likely reflects the interaction between mechanical load and the host's capacity to withstand it.[4,7]

Accordingly, this study had two aims: to determine whether intraoperative blood loss independently predicts pedicle screw loosening after controlling for surgical complexity and biomechanical–biological factors, and to compare the predictive performance of composite indices representing these domains. We hypothesized that the association between EBL and loosening would diminish after multivariate adjustment and that the biomechanical–biological composite would demonstrate superior discriminative ability.

MATERIALS and METHODS

Study Design and Patient Selection

This retrospective matched cohort study was conducted at a tertiary spine center between January 1, 2020, and November 15, 2025. Institutional ethics approval was obtained from the Kanuni Sultan Süleyman Training and Research Hospital Ethics Committee (Approval No. KAEK/2026.03.110; Approval Date: March 11, 2026) in accordance with the Declaration of Helsinki. The requirement for informed consent was waived due to the retrospective study design. A total of 434 consecutive patients undergoing posterior lumbar instrumented fusion for degenerative pathology were screened. Inclusion criteria were age ≥18 years, available preoperative MRI and standing radiographs, and adequate clinical/imaging data. Exclusion criteria included infection, tumor, congenital deformity, inadequate imaging, prior hip or pelvic fracture, bilateral hip arthroplasty, and systemic inflammatory spinal disease.

After application of the inclusion and exclusion criteria, 220 eligible patients remained, including 80 patients with revision-confirmed pedicle screw loosening and 140 potential controls without loosening. Controls were selected using 1:1 nearest-neighbor matching by age (±3 years) and sex, without replacement. The final matched cohort comprised 160 patients (80 cases and 80 matched controls).

Definition of Pedicle Screw Loosening

Pedicle screw loosening was defined based on radiographic findings and the clinical indication for revision surgery, in accordance with previously established criteria.[2,4] Radiographic loosening was identified as the presence of a radiolucent zone of at least 1 mm surrounding the pedicle screw ("halo sign") on standard radiographs, which is a commonly accepted diagnostic threshold in the literature.[2,3]

To enhance diagnostic specificity, isolated or partial radiolucency without circumferential involvement was not considered sufficient. Only cases in which radiographic loosening was accompanied by persistent clinical symptoms and resulted in revision surgery were classified as clinically significant screw loosening.[4]

Patients in the control group were defined as those without radiographic evidence of screw loosening and without revision surgery during the follow-up period.

Imaging Protocols and Radiological Assessment

MRI Acquisition and Paraspinal Muscle Analysis

All preoperative magnetic resonance imaging (MRI) examinations were performed using a 1.5-T system. Axial T2-weighted images were used for the evaluation of paraspinal muscle morphology. To ensure consistency and minimize inter-level variability, all measurements were standardized at the L4–L5 intervertebral disc level.[7,10]

Paraspinal muscles, including the multifidus, erector spinae, and quadratus lumborum, were analyzed. Cross-sectional area (CSA) was measured by manually tracing the fascial boundaries of each muscle using region-of-interest (ROI) analysis. Fat infiltration (FI) was assessed semi-quantitatively within the same ROI using a threshold-based segmentation method.[7,10] During the thresholding process, reference signal characteristics were obtained from three separate regions of interest placed within the subcutaneous fat on the same axial slice, and the mean value of these measurements was used as the reference standard. Based on this reference, fat and muscle tissues were differentiated, and FI was calculated as the percentage of the fat signal area relative to the total muscle area.[10]

All measurements were performed independently by two observers blinded to clinical and radiological outcomes. The mean values of the two measurements were used for analysis. Interobserver reliability was assessed using intraclass correlation coefficients (ICCs).[10]

Radiographic Assessment of Spinopelvic Parameters

Spinopelvic parameters were assessed using standing radiographs obtained as part of routine clinical evaluation. When available, full-length standing scoliosis radiographs were preferred; otherwise, standing lateral lumbosacral radiographs were used for analysis.[6,11]

The following parameters were measured: pelvic incidence (PI), pelvic tilt (PT), sacral slope (SS), lumbar lordosis (LL), and PI–LL mismatch.[6]

All measurements were performed using the institutional picture archiving and communication system (PACS) with integrated digital measurement tools. Evaluations were conducted independently by two observers who were blinded to clinical outcomes. In cases of discrepancies greater than 5°, measurements were re-evaluated, and a consensus value was obtained.

Interobserver reliability for all spinopelvic parameters was assessed using intraclass correlation coefficients (ICCs).[11]

Pelvic incidence was defined as the angle between the perpendicular to the sacral endplate at its midpoint and the line connecting this point to the femoral head axis. Pelvic tilt was measured as the angle between the vertical and the line connecting the midpoint of the sacral endplate to the femoral head axis. Sacral slope was defined as the angle between the sacral endplate and the horizontal plane. Lumbar lordosis was measured between the superior endplate of L1 and the superior endplate of S1.[6]

PI–LL mismatch was calculated as the difference between pelvic incidence and lumbar lordosis and was used as the primary parameter representing sagittal alignment.[6,11]

CT Evaluation of Screw Loosening

Pedicle screw loosening was evaluated using postoperative computed tomography (CT) imaging. Loosening was defined as the presence of a radiolucent zone of at least 1 mm surrounding the pedicle screw, consistent with previously established radiographic criteria.[2,3,11]

Only cases demonstrating circumferential radiolucency around the screw were considered positive for loosening, while focal or equivocal radiolucent changes were not accepted as diagnostic.[3,11]

To ensure clinical relevance, only patients with CT-confirmed screw loosening who underwent revision surgery due to persistent symptoms were included in the loosening group.[4]

CT images were independently evaluated by two observers blinded to clinical data. In cases of disagreement, a consensus decision was reached. Interobserver reliability was assessed using intraclass correlation coefficients (ICCs).[11]

Patients in the control group were defined as those without radiographic evidence of screw loosening on available imaging and without revision surgery during the follow-up period.

Bone Mineral Density Assessment (DEXA)

Bone mineral density (BMD) was assessed using dual-energy X-ray absorptiometry (DEXA) as part of routine preoperative evaluation. Measurements were obtained from the lumbar spine and/or femoral neck regions, depending on availability.

T-scores were recorded for each patient and used for analysis. In cases where measurements from multiple anatomical regions were available, the lowest T-score was selected to represent the patient's bone quality.

Based on World Health Organization (WHO) criteria, osteoporosis was defined as a T-score ≤-2.5, osteopenia as a T-score between -1.0 and -2.5, and normal bone density as a T-score ≥-1.0.[12]

Surgical Data and Perioperative Parameters

Surgical and perioperative data were retrospectively obtained from operative records and hospital systems, including estimated blood loss (EBL), operative time, number of instrumented levels, and lumbosacral involvement. EBL was recorded in milliliters from anesthesiology records based on suction volume and intraoperative estimation. Operative time was defined as the interval from skin incision to closure. Instrumented levels were determined by the extent of posterior lumbar fixation, and lumbosacral involvement was defined as the inclusion of the L5–S1 segment.

Matching Procedure

A matched cohort design was used to minimize confounding. Patients with pedicle screw loosening were matched 1:1 with controls without loosening using a nearest-neighbor approach based on age and sex.[13] Sex was matched exactly, and controls were selected based on the closest age, with a caliper of ±3 years when possible. Matching was performed without replacement.

Controls were drawn from the same institutional cohort and met identical inclusion and exclusion criteria, except for the absence of screw loosening. The final cohort included 160 patients: 80 with loosening and 80 matched controls.

Composite Index Construction

To evaluate the relative contributions of surgical burden and underlying biomechanical–biological vulnerability, composite indices were constructed using standardized variables.

All continuous variables included in the composite indices were transformed into z-scores using the formula (value – mean)/standard deviation to ensure comparability between parameters with different units and scales.

Surgical Complexity Index

The surgical complexity index was designed to represent the overall procedural burden. This index was calculated by summing the standardized values of estimated blood loss (EBL), number of instrumented levels, and operative time.

Higher values of the surgical complexity index indicated greater surgical burden.

Biomechanical–Biological Index

The biomechanical–biological index was constructed to reflect the intrinsic vulnerability of the spinal environment. This index included standardized values of PI–LL mismatch, multifidus fat infiltration (FI), and bone mineral density as represented by DEXA T-score.

Since lower T-scores indicate poorer bone quality, the T-score variable was inversely coded prior to standardization so that higher index values consistently represented greater biomechanical–biological risk.

Statistical Analysis

All analyses were performed using SPSS (Version 25.0; IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean±standard deviation or median (minimum–maximum) according to distribution, and categorical variables as frequencies and percentages. Normality was assessed with the Shapiro–Wilk test. Between-group comparisons used the independent-samples t-test or Mann–Whitney U test for continuous variables and the chi-square or Fisher's exact test for categorical variables, as appropriate.

Spearman correlation analysis assessed associations between variables. Univariate logistic regression identified factors associated with pedicle screw loosening; variables meeting a predefined threshold of p=0.10 were entered into the multivariate model. Multicollinearity was evaluated using variance inflation factors (VIFs), with <3 considered acceptable.

Receiver operating characteristic (ROC) analysis evaluated discriminative performance, with the area under the curve (AUC) calculated and compared using DeLong's method.[14] Model calibration was assessed using the Hosmer–Lemeshow goodness-of-fit test and the Brier score. Interobserver reliability for radiological measurements was assessed using intraclass correlation coefficients (ICCs).[15]

Statistical significance was defined at a p=0.05 threshold.

Table 1. Baseline characteristics of the matched cohort

Variable	Loosening (n=80)	Control (n=80)	p
Age (years)	63.6±8.0	63.3±7.7	0.779
BMI (kg/m²)	28.3±5.1	30.8±4.8	0.002
Operative time (min)	199.8±55.1	162.5±57.1	0.001
EBL (mL)	712.9±269.6	456.2±193.6	0.001
Instrumented levels	3.8±1.6	2.5±0.9	0.001
PI (°)	52.6±10.9	54.3±10.3	0.316
Lumbar lordosis (°)	38.4±12.5	52.9±14.1	0.001
PI–LL mismatch (°)	14.3±10.2	1.4±8.3	0.001
Multifidus FI (%)	28±8.0	21±5.5	0.001
DEXA T-score	−2.2±0.9	−0.7±1.4	0.001

Values are mean±standard deviation. Statistical significance was defined at a p=0.05 threshold. BMI: Body mass index; EBL: Estimated blood loss; PI: Pelvic incidence; LL: Lumbar lordosis; FI: Fat infiltration; DEXA: Dual-energy X-ray absorptiometry

RESULTS

Patient Cohort and Matching

A total of 434 patients who underwent posterior lumbar instrumented fusion during the study period were screened for eligibility. After applying the inclusion and exclusion criteria, a matched cohort was constructed using a 1:1 nearest-neighbor approach based on age and sex, resulting in a final study population of 160 patients, including 80 patients with pedicle screw loosening and 80 matched controls.

Following matching, there were no significant differences between the groups in terms of age or sex distribution, confirming adequate demographic balance between the two groups.

Demographic and Clinical Characteristics

Baseline demographic and clinical characteristics of the matched cohort are presented in Table 1. There were no significant differences between the groups in terms of age or sex distribution. Body mass index (BMI) was lower in the loosening group compared with the control group (p=0.010). No significant differences were observed between the groups in terms of other baseline clinical characteristics (Fig. 1).

Surgical Parameters

Surgical and perioperative parameters are summarized in Table 1. Operative time was longer in the loosening group compared with controls, and estimated blood loss (EBL) was higher in patients with screw loosening (Fig. 2). The number of instrumented levels was greater in the loosening group.

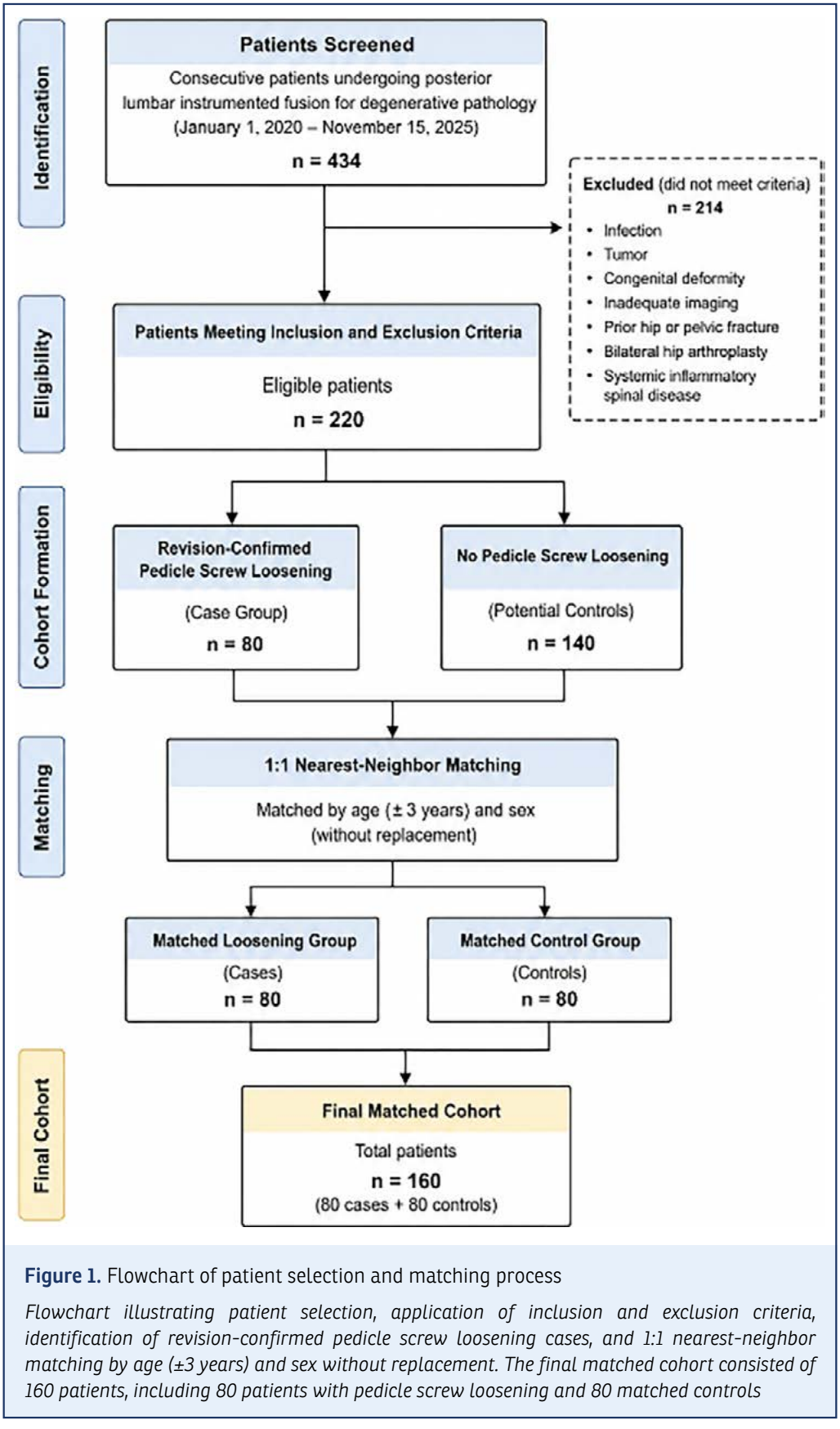


Figure 1. Flowchart of patient selection and matching process

Flowchart illustrating patient selection, application of inclusion and exclusion criteria, identification of revision-confirmed pedicle screw loosening cases, and 1:1 nearest-neighbor matching by age (±3 years) and sex without replacement. The final matched cohort consisted of 160 patients, including 80 patients with pedicle screw loosening and 80 matched controls

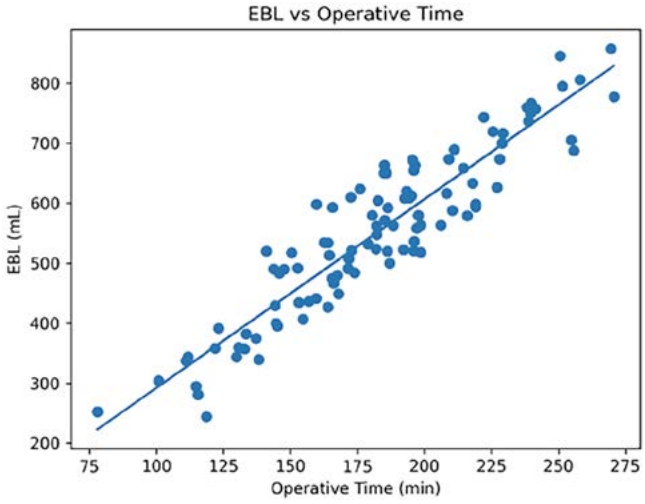


Figure 2. Relationship between estimated blood loss and surgical complexity parameters

Scatter plots showing the relationship between estimated blood loss (EBL) and surgical complexity parameters. EBL was positively correlated with operative time, indicating that intraoperative blood loss reflects procedural burden rather than an independent determinant of pedicle screw loosening. The solid line represents the linear regression fit

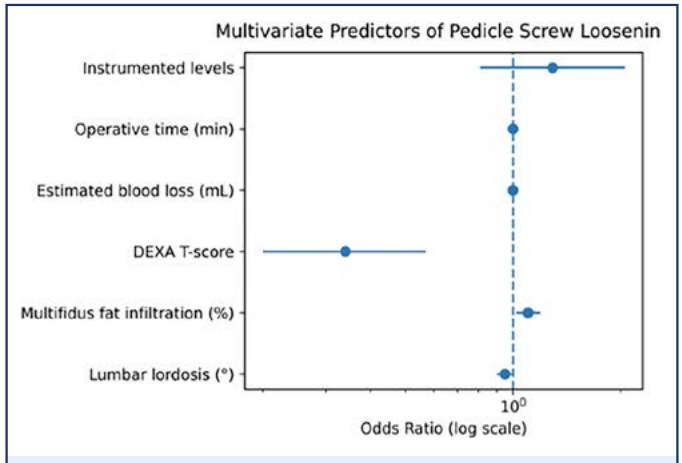


Figure 3. Multivariate predictors of pedicle screw loosening

Forest plot of the multivariate conditional logistic regression analysis. Odds ratios (ORs) with 95% confidence intervals (CIs) are presented. Lumbar lordosis, multifidus fat infiltration, and DEXA T-score were independently associated with pedicle screw loosening, whereas estimated blood loss, operative time, and instrumented levels were not significant after adjustment. The vertical dashed line indicates an OR of 1

Radiographic and Paraspinal Muscle Findings

Radiographic and paraspinal muscle parameters are presented in Table 1. Pelvic incidence did not differ significantly between the groups. Lumbar lordosis was significantly lower in the loosening group, resulting in a greater PI–LL mismatch. Multifidus fat infiltration was higher in the loosening group compared with controls.

Logistic Regression Analysis

Conditional logistic regression analysis was performed to account for the matched study design.

In the adjusted model, lumbar lordosis (OR=0.95, 95% CI: 0.90–0.99, p=0.031), multifidus fat infiltration (OR=1.10, 95% CI: 1.02–1.19, p=0.009), and DEXA T-score (OR=0.34, 95% CI: 0.20–0.57, p=0.001) remained independently associated with pedicle screw loosening (Table 2). These findings are illustrated in Figure 3.

Surgical parameters, including estimated blood loss, operative time, and number of instrumented levels, were not independently associated with screw loosening after adjustment.

Interobserver Reliability

Interobserver reliability was high for all measurements. Intraclass correlation coefficient values were 0.90 for spinopelvic parameters and 0.88 for paraspinal muscle measurements.

Composite Index Analysis

Receiver operating characteristic (ROC) analysis evaluated the discriminative performance of the composite indices. The surgical complexity index showed moderate predictive ability, whereas the biomechanical–biological index reflect-

Table 2. Conditional logistic regression analysis of factors associated with pedicle screw loosening

Variable	OR (95% CI)	p
BMI	1.08 (0.96–1.23)	0.180
Operative time (min)	1.00 (0.99–1.01)	0.880
EBL (mL)	1.00 (0.998–1.003)	0.640
Instrumented levels	1.29 (0.81–2.05)	0.280
Lumbar lordosis (°)	0.95 (0.90–0.99)	0.031
PI–LL mismatch (°)	1.05 (0.97–1.14)	0.210
Multifidus FI (%)	1.10 (1.02–1.19)	0.009
DEXA T-score	0.34 (0.20–0.57)	0.001

Odds ratios (ORs) and 95% confidence intervals (CIs) were estimated using conditional logistic regression to account for the matched design. Variables meeting a predefined threshold of p=0.10 in univariate analysis were included in the multivariate model. Statistical significance was defined at a p=0.05 threshold. OR: Odds ratio; CI: Confidence interval; BMI: Body mass index; EBL: Estimated blood loss; FI: Fat infiltration; DEXA: Dual-energy X-ray absorptiometry

Table 3. Predictive performance of composite indices for pedicle screw loosening

Model	AUC	95% CI	z-value	p
Surgical complexity index	0.771	0.693–0.849	–	0.001
Biomechanical–biological index	0.908	0.861–0.955	–	0.001
Comparison (DeLong test)	–	–	3.91	0.001

Discriminative performance was evaluated using receiver operating characteristic (ROC) analysis. Statistical significance was defined at a p=0.05 threshold. AUC: Area under the curve; CI: Confidence interval

ed higher performance (Table 3, Fig. 4). Comparative analysis confirmed the superior predictive accuracy of the biomechanical–biological index.

Calibration Analysis

The final multivariable model demonstrated acceptable calibration, with a Hosmer–Lemeshow goodness-of-fit statistic of χ^2=14.74 (df=8, p=0.064). The Brier score was 0.135. The expanded model, which additionally included estimated blood loss, operative time, and number of instrumented levels, also demonstrated acceptable calibration (Hosmer–Lemeshow χ^2=8.37, df=8, p=0.398), with a Brier score of 0.123. The area under the receiver operating characteristic curve was 0.883 for the final model and 0.899 for the expanded model.

Sensitivity Analysis

Sensitivity analyses were performed to assess robustness. Influential observations were examined using Cook's distance, and models were re-estimated after excluding high-leverage cases, with no change in the associations between key predictors and outcomes. Robust regression using heteroskedasticity-consistent standard errors (HC3) showed no meaningful change in statistical significance or effect size. Nonparametric correlation analyses (Spearman) confirmed a consistent direction and strength of associations. Reduced models including only core predictors were comparable to full models, with no substantial differences in performance or variable significance. These findings indicate that the results are stable.

DISCUSSION

The present study suggests that pedicle screw loosening is primarily associated with intrinsic biomechanical and biological factors rather than surgical complexity. Although intraoperative parameters such as estimated blood loss (EBL), operative time, and instrumented levels were associated with loosening in univariate analyses, these associations did not persist after conditional logistic regression. In contrast, lumbar lordosis, multifidus fat infiltration, and bone mineral density remained independently associated with pedicle screw loosening, underscoring the role of sagittal alignment and tissue quality.[6,7,12]

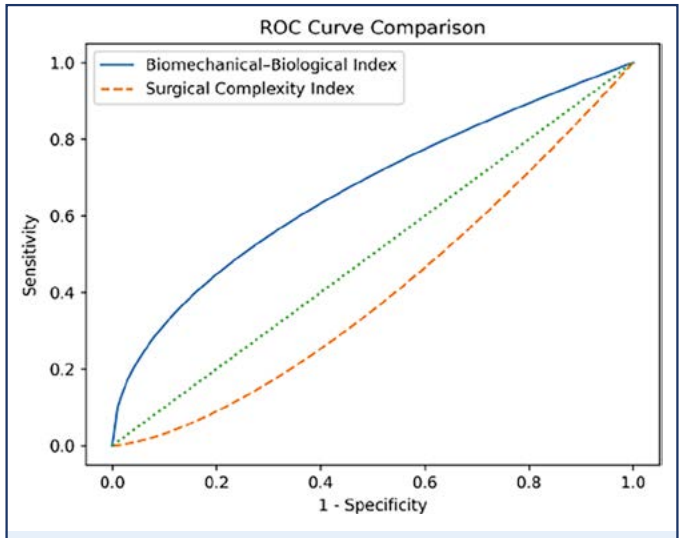


Figure 4. Receiver operating characteristic analysis of composite indices for predicting pedicle screw loosening

Receiver operating characteristic (ROC) curves comparing the surgical complexity index and the biomechanical–biological index. The biomechanical–biological index showed higher discriminative ability, reflected by a greater area under the curve (AUC). Model performance differed significantly according to DeLong's test. The diagonal dashed line indicates no discrimination (AUC=0.5)

A key strength is the matched cohort design with conditional logistic regression,[13] minimizing demographic confounding by accounting for age, sex, and the matched structure. This approach improves the estimation of independent effects, and the persistence of biomechanical and biological predictors supports robustness.

Surgical complexity requires careful interpretation. EBL was higher in the loosening group (≈250 mL difference), but its clinical relevance appears limited. It did not retain independent predictive value after adjustment, indicating that it reflects procedural burden rather than a direct biological effect.[8,9,16] Operative time and instrumented levels showed similar patterns. These findings suggest that more extensive procedures are indirectly associated with loosening through underlying biomechanical and biological vulnerability.[8,16] Sensitivity analyses confirmed consistency across models and the absence of outlier- or assumption-driven effects.

Sagittal alignment emerged as a key determinant. Reduced lumbar lordosis was independently associated with pedicle screw loosening, supporting the importance of restoring sagittal balance.[6,11,17] PI–LL mismatch was associated with loosening in univariate analysis but not after adjustment, likely due to collinearity, indicating lumbar lordosis as the more direct clinical parameter.

Paraspinal muscle quality, particularly multifidus fat infiltration, was another factor independently associated with pedicle screw loosening. Given its role in segmental stabilization, degeneration may increase micromotion at the bone–implant interface,[7,10,18] consistent with prior evidence.

Bone mineral density was the strongest predictor; lower DEXA T-scores were associated with an increased risk of loosening, highlighting the importance of bone quality and the need for preoperative optimization.[5,12,19]

Composite index analysis supported these findings. The surgical complexity index showed moderate performance, whereas the biomechanical–biological index reflected higher discriminative ability and outperformed the surgical model, indicating a greater contribution of biomechanical and biological factors. Integration of alignment, muscle quality, and bone density may provide a more meaningful framework for risk stratification.

Clinically, these results support prioritizing preoperative optimization and biomechanical planning. Restoration of sagittal alignment, preservation of paraspinal muscle integrity, and management of osteoporosis appear central, whereas surgical technique alone may be insufficient under unfavorable biological conditions.

Limitations

This study is limited by its retrospective design and potential selection bias, although matching and conditional regression partially mitigate these limitations.[13] Matching was restricted to age and sex, and residual confounding cannot be excluded. The single-center design may limit generalizability. Fat infiltration was assessed using a semi-quantitative MRI-based method rather than advanced quantitative techniques.[10,20] Loosening was defined by revision-confirmed cases, increasing specificity but potentially underestimating subclinical cases.[4,20]

CONCLUSION

In this retrospective matched cohort study, pedicle screw loosening was primarily associated with biomechanical and biological factors rather than surgical complexity.[6,7,12] Lumbar lordosis, paraspinal muscle degeneration, and bone mineral density were independently associated with pedicle screw loosening, whereas intraoperative parameters such as blood loss did not retain independent predictive value after multivariable adjustment and instead reflected procedural burden.[8,9] Composite modeling demonstrated superior predictive performance of biomechanical–biological factors, supporting their use in risk assessment and clinical decision-making.

Disclosures

Ethics Committee Approval: The study was approved by the Kanuni Sultan Süleyman Training and Research Hospital Ethics Committee (No: KAEK/2026.03.110, Date: 11/03/2026).

Informed Consent: Informed consent was waived due to the retrospective nature of the study.

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

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

Use of AI for Writing Assistance: No artificial intelligence–assisted technology was used in the design, analysis, or writing of this study.

Author Contributions: Concept – E.S.; Design – E.S.; Supervision – E.S.; Data collection and/or processing – E.S.; Data analysis and/or interpretation –E.S.; Literature search – E.S.; Writing – E.S.; Critical review – L.A., H.D.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Kim HJ, Lenke LG, Shaffrey CI, Van Alstyne EM, Skelly AC. Proximal junctional kyphosis as a distinct form of adjacent segment pathology after spinal deformity surgery: a systematic review. Spine (Phila Pa 1976) 2012;37(22 Suppl):S144–64. [CrossRef]
2. Bokov A, Bulkin A, Aleynik A, Kutlaeva M, Mlyavykh S. Pedicle screws loosening in patients with degenerative diseases of the lumbar spine: potential risk factors and relative contribution. Global Spine J 2019;9:55–61. [CrossRef]
3. Bredow J, Boese CK, Werner CML, Siewe J, Löhrer L, Zarghooni K, et al. Predictive validity of preoperative CT scans and the risk of pedicle screw loosening in spinal surgery. Arch Orthop Trauma Surg 2016;136:1063–7. [CrossRef]
4. Galbusera F, Volkheimer D, Reitmaier S, Berger-Roscher N, Kienle A, Wilke HJ. Pedicle screw loosening: a clinically relevant complication? Eur Spine J 2015;24:1005–16. [CrossRef]
5. Schreiber JJ, Anderson PA, Rosas HG, Buchholz AL, Au AG. Hounsfield units for assessing bone mineral density and strength: a tool for osteoporosis management. J Bone Joint Surg Am 2011;93(11):1057–63. [CrossRef]
6. Schwab F, Lafage V, Patel A, Farcy JP. Sagittal plane considerations and the pelvis in the adult patient. Spine (Phila Pa 1976) 2009;34:1828–33. [CrossRef]
7. Fortin M, Macedo LG. Multifidus and paraspinal muscle group cross-sectional areas of patients with low back pain and control patients: a systematic review with a focus on blinding. Phys Ther 2013;93:873–88. [CrossRef]

8. Kim JB, Park SW, Lee YS, Nam TK, Park YS, Kim YB. The effects of spinopelvic parameters and paraspinal muscle degeneration on S1 screw loosening. J Korean Neurosurg Soc 2015;58:357–62. [CrossRef]
9. Raman T, Vasquez-Montes D, Varlotta C, Passias PG, Errico TJ. Decision tree-based modelling for identification of predictors of blood loss and transfusion requirement after adult spinal deformity surgery. Int J Spine Surg 2020;14:87–95. [CrossRef]
10. Crawford RJ, Filli L, Elliott JM, Nanz D, Fischer MA, Marcon M, et al. Age- and level-dependence of fatty infiltration in lumbar paravertebral muscles of healthy volunteers. AJNR Am J Neuroradiol 2016;37:742–8. [CrossRef]
11. Le Huec JC, Thompson W, Mohsinaly Y, Barrey C, Faundez A. Sagittal balance of the spine. Eur Spine J 2019;28:1889–905. [CrossRef]
12. Kanis JA, Harvey NC, McCloskey E, Bruyère O, Veronese N, Lorentzon M, et al. Algorithm for the management of patients at low, high and very high risk of osteoporotic fractures. Osteoporos Int 2020;31:1–12. [CrossRef]
13. Benedetto U, Head SJ, Angelini GD, Blackstone EH. Statistical primer: propensity score matching and its alternatives. Eur J Cardiothorac Surg 2018;53:1112–7. [CrossRef]
14. Sun X, Xu W. Fast implementation of DeLong's algorithm for comparing the areas under correlated receiver operating characteristic curves. IEEE Signal Process Lett 2014;21:1389–93. [CrossRef]
15. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. J Chiropr Med 2016;15:155–63. [CrossRef]
16. Schwab FJ, Hawkinson N, Lafage V, Smith JS, Hart R, Mundis G, et al. Risk factors for major peri-operative complications in adult spinal deformity surgery: a multi-center review of 953 consecutive patients. Eur Spine J 2012;21:2603–10. [CrossRef]
17. Lafage R, Schwab F, Glassman S, Bess S, Harris B, Scheer JK, et al. Age-adjusted alignment goals have the potential to reduce PJK. Spine (Phila Pa 1976) 2017;42:1275–82. [CrossRef]
18. Crawford RJ, Elliott JM, Volken T. Change in fatty infiltration of lumbar multifidus, erector spinae, and psoas muscles in asymptomatic adults of Asian or Caucasian ethnicities. Eur Spine J 2017;26:3059–67. [CrossRef]
19. Zou D, Sun Z, Zhou S, Zhong W, Li W. Hounsfield units value is a better predictor of pedicle screw loosening than the T-score of DXA in patients with lumbar degenerative diseases. Eur Spine J 2020;29:1105–11. [CrossRef]
20. Hu ZJ, He J, Zhao FD, Fang XQ, Zhou LN, Fan SW. An assessment of the intra- and inter-reliability of the lumbar paraspinal muscle parameters using CT scan and magnetic resonance imaging. Spine (Phila Pa 1976). 2011;36:E868–74. [CrossRef]

DOI: 10.14744/cm.2026.57855
Compreh Med 2026;18(Suppl. 1):47-52

RESEARCH ARTICLE

Frequency of Sarcopenia and Associated Factors in Older Adults with Non-dialysis Dependent Chronic Kidney Disease: A Cross-sectional Study

Tuğçe Emiroğlu Gedik[1], Onur Kağan Öztaş[2], Mustafa Ortaboz[2], Aylin Şendoğan[3], Murat Akarsu[2]

[1]Division of Geriatrics, Department of Internal Medicine, İstanbul University-Cerrahpaşa, Cerrahpaşa Faculty of Medicine, İstanbul, Türkiye
[2]Department of Internal Medicine, University of Health Sciences, İstanbul Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
[3]Department of Nutrition and Dietetics, University of Health Sciences, İstanbul Kanuni Sultan Süleyman Research and Training Hospital, İstanbul, Türkiye

ABSTRACT

Objective: This cross-sectional study aimed to determine the prevalence of sarcopenia and investigate the associated clinical and biochemical factors in geriatric patients with non-dialysis chronic kidney disease (CKD).

Materials and Methods: Seventy-one pre-dialysis CKD patients followed in geriatric and nephrology outpatient clinics were enrolled. Sarcopenia was diagnosed according to the EWGSOP2 criteria using the skeletal muscle index (SMI; bioelectrical impedance analysis), handgrip strength (dynamometer), 4-metre gait speed, and five-times sit-to-stand test. Patients with probable sarcopenia were included in the sarcopenic group. The Mann-Whitney U test, Fisher's exact test, and univariate logistic regression analysis were applied.

Results: The median age was 71.8±4.4 years, and 45.1% of patients were male. The prevalence of sarcopenia was 22.5% (n=16), with defined sarcopenia at 14.1% and probable sarcopenia at 8.5%. BMI was significantly lower in patients with sarcopenia (26.2 vs. 31.9 kg/m^2; p=0.009). The SMI, fat-free mass index (FFMI), and handgrip strength were significantly lower in the sarcopenic group (all p<0.05). Hemoglobin showed a strong positive correlation with SMI (r=+0.524; p<0.001) and handgrip strength (r=+0.437; p<0.001). In univariate logistic regression analysis, BMI was the only significant independent predictor (OR=0.801; 95% CI: 0.677–0.947; p=0.009).

Conclusion: Sarcopenia prevalence was 22.5% in non-dialysis CKD patients, and low BMI was an independent risk indicator. The strong correlation between hemoglobin and muscle parameters supports the role of renal anemia in the pathophysiology of sarcopenia. Early sarcopenia screening and body composition assessment are recommended for CKD follow-up.

Keywords: Anemia, chronic kidney disease, older adults, sarcopenia

How to cite this article: Emiroğlu Gedik T, Öztaş OK, Ortaboz M, Şendoğan A, Akarsu M. Frequency of Sarcopenia and Associated Factors in Older Adults with Non-dialysis Dependent Chronic Kidney Disease: A Cross-sectional Study. Compreh Med 2026;18(Suppl. 1):47-52.

INTRODUCTION

Sarcopenia, a multifactorial clinical syndrome, is characterised by progressive loss of skeletal muscle mass, muscle strength, and physical performance. It is associated with loss of functionality, increased risk of falls, hospitalisations, morbidity, and mortality. The 2018 updated consensus statement of the European Working Group on Sarcopenia in Older People (EWGSOP2) identifies low muscle strength as the primary diagnostic criterion; confirmation of low muscle mass is additionally required for a diagnosis of "sarcopenia," while impairment in all three components defines "severe sarcopenia." Patients with isolated reduced muscle strength are classified as having "probable sarcopenia," representing an important window of opportunity for intervention.[1,2]

Chronic kidney disease (CKD) has become a major public health problem globally, largely attributable to demographic ageing and the increasing incidence of hypertension and diabetes. Population-based studies indicate a CKD prevalence of

Address for Correspondence: Tuğce Emiroğlu Gedik, Division of Geriatrics, Department of Internal Medicine, İstanbul University-Cerrahpaşa, Cerrahpaşa Faculty of Medicine, İstanbul, Türkiye
E-mail: tugceemiroglu@yahoo.com.tr **ORCID ID:** 0000-0002-5550-6477

Received date: 19.06.2026
Revised date: 05.08.2026
Accepted date: 12.08.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

7.2% among individuals aged 30 years and older, with this figure rising to between 23.4% and 35.8% in older adults.[3,4] In Türkiye, the CREDIT study identified CKD rates of 32.7% in the 60–69 age group, 41.3% in the 70–79 age group, and 54.7% in individuals over 80 years of age.[5] The pathophysiology of CKD contributes to sarcopenia through several converging mechanisms, including the accumulation of uraemic toxins, metabolic acidosis that accelerates muscle proteolysis, disorders in insulin and IGF-1 signalling leading to anabolic resistance, vitamin D deficiency with direct myotoxic effects, and systemic inflammation.[6,7] The independent roles of CKD-specific factors—such as proteinuria, renal anaemia, and secondary hyperparathyroidism—in the development and progression of sarcopenia remain incompletely understood.[8]

Comprehensive sarcopenia assessments using the EWGSOP2 criteria in predialysis CKD patients remain limited. Using varying definitions, the prevalence of sarcopenia ranges from 5.9% to 14% in patients with CKD not receiving renal replacement therapy and may reach up to 42.2% in haemodialysis patients.[9–14] Currently, there is a paucity of population-specific data from Türkiye employing the EWGSOP2 framework in predialysis CKD patients. Utilising the EWGSOP2 criteria, the present study investigated the prevalence of sarcopenia among predialysis CKD patients and examined associated clinical and biochemical factors, with particular emphasis on eGFR staging, type 2 diabetes mellitus (T2DM) aetiology, haemoglobin concentrations, and proteinuria.

MATERIALS and METHODS

Study Design and Patient Selection

This cross-sectional study was conducted at the Geriatrics and Nephrology Outpatient Clinics of İstanbul Kanuni Sultan Süleyman Training and Research Hospital from August 1, 2023, to November 1, 2023. Participants eligible for inclusion were individuals aged 65 years and older who had been under follow-up for a minimum of six months with a diagnosis of chronic kidney disease (CKD) and had provided written informed consent. The exclusion criteria included active dialysis treatment, a history of kidney transplantation, active malignancy, decompensated heart failure (NYHA Stage III–IV), active infection, advanced liver cirrhosis, and orthopaedic or neuromuscular disorders that precluded bilateral handgrip measurement. Of the 81 patients initially assessed, five declined participation, four were excluded due to significant oedema, and one was excluded due to missing data, resulting in the inclusion of 71 patients. The estimated glomerular filtration rate (eGFR) was calculated using the 2009 CKD-EPI creatinine equation, and chronic kidney disease (CKD) staging was conducted in accordance with the KDIGO 2012 guideline.[15,16] Approval from the Ethics Committee of the University of Health Sciences Kanuni Sultan Suleyman Training and Research Hospital was secured for this study (Decision No: 2023.07.89; Decision date: 13.07.2023). Furthermore, the consent process was conducted in writing for all participants, conforming to the Declaration of Helsinki guidelines.

Definitons of Sarcopenia

Body composition was assessed using an eight-electrode multifrequency bioelectrical impedance analysis (BIA) device (InBody 270, InBody Co., Ltd., Seoul, South Korea). Measurements were performed in the morning after at least eight hours of fasting; participants abstained from vigorous physical activity and alcohol for 24 hours prior to measurement. All measurements were conducted with participants wearing light clothing, without shoes or socks, standing on the foot electrodes and gripping the hand electrodes. Parameters recorded included the appendicular skeletal muscle index (ASMI), calculated as appendicular skeletal muscle mass (kg) divided by height squared (m^2); total body fat mass (kg); percent body fat; and BMI. Cut-off values for low appendicular muscle mass were ASMI<7.0 kg/m^2 for men and <5.5 kg/m^2 for women.[1]

Handgrip strength was measured on the dominant hand using a hydraulic hand dynamometer (Baseline Hydraulic Hand Dynamometer 12–024, Fabrication Enterprises, New York, USA); the mean of three measurements was recorded. Thresholds for low handgrip strength were <27 kg for men and <16 kg for women. For the five-times chair stand test, a time of ≥15 seconds indicated low muscle strength.[1] Gait speed was assessed over 4 metres; a value <0.8 m/s defined poor physical performance. Patients with isolated low muscle strength were classified as having probable sarcopenia; those with additional low ASMI were classified as having sarcopenia; and those with impairment in all three components were classified as having severe sarcopenia. Probable sarcopenia cases were included in the sarcopenic group for all analyses.

Clinical and Biochemical Assessment

Serum creatinine, BUN, albumin, hemoglobin (Hb), HbA1c, CRP, PTH, calcium, phosphorus, sodium, and potassium were measured from fasting morning blood samples. Proteinuria was assessed using a 24-hour urine collection. Body weight, height, and BMI were measured under standardised conditions.

Statistical Analysis

The Shapiro-Wilk test was employed to assess the normality of continuous variables. Variables with a normal distribution were expressed as mean±standard deviation (mean±SD),

Table 1. General characteristics and between-group comparison of patients

Variable	All patients (n=71)	Sarcopenic (n=16)	Non-sarcopenic (n=55)	p
Demographic characteristics				
Age (years)	71.8±4.4	71.2±3.9	71.9±4.5	0.588
Sex, male n (%)	32 (45.1)	4 (25.0)	28 (50.9)	0.089
BMI (kg/m²)	31.1 (27.8–34.5)	26.2 (23.6–30.8)	31.9 (28.7–35.0)	**0.009**
Renal function parameters				
eGFR (mL/min/1.73 m²)	42.1 (32.3–50.4)	37.9 (27.8–49.3)	43.2 (34.8–50.4)	0.285
eGFR <30, n (%)	16 (22.5)	6 (37.5)	10 (18.2)	0.170
Creatinine (mg/dL)	1.39 (1.13–1.83)	1.47 (1.20–1.89)	1.39 (1.13–1.83)	0.571
BUN (mg/dL)	24.7 (19.8–32.2)	28.5 (21.3–33.4)	24.7 (19.7–32.2)	0.359
Proteinuria (mg/day)	130.0 (60.0–450.0)	70.0 (40.0–155.0)	165.0 (80.0–525.0)	0.055
Biochemical parameters				
Haemoglobin (g/dL)	12.6 (11.5–13.7)	11.9 (11.4–12.6)	12.7 (11.6–13.9)	0.155
Albumin (g/L)	44.0 (42.0–46.0)	43.0 (42.0–44.0)	44.0 (42.0–46.0)	0.247
CRP (mg/L)	2.8 (1.1–6.2)	2.4 (1.0–7.8)	2.8 (1.2–6.1)	0.865
PTH (pg/mL)	64.0 (45.0–91.0)	80.5 (57.0–90.0)	62.0 (45.0–91.0)	0.397
Vitamin D (mg/L)	21.19.7	20,67.1	21.210.3	0.845
Phosphorus (mg/dL)	3.5 (3.2–4.0)	3.8 (3.5–4.4)	3.5 (3.2–4.0)	0.114
Potassium (mEq/L)	4.8 (4.4–5.0)	4.9 (4.6–5.1)	4.8 (4.4–5.0)	0.323
Body composition and performance				
ASMI (kg/m²)	7.3 (6.6–8.4)	5.9 (5.7–6.8)	8.0 (6.9–8.7)	**<0.001**
FFMI (kg/m²)	18.6 (17.3–20.7)	16.2 (15.7–16.6)	19.2 (17.7–20.8)	**<0.001**
Handgrip strength (kg)	25.0 (18.0–32.0)	16.0 (15.0–24.5)	28.0 (22.0–37.8)	**0.002**
Gait speed (m/s)	0.75 (0.67–0.92)	0.75 (0.63–0.83)	0.75 (0.67–0.92)	0.263
Chair stand time (s)	12.0 (10.0–14.0)	14.0 (11.2–16.0)	12.0 (10.0–13.8)	0.053
T2DM (aetiology), n (%)	35 (49.3)	11 (68.8)	24 (43.6)	0.094

Data are presented as mean±SD (normally distributed) and median (IQR) or n (%). Bold P-values indicate statistical significance (P<0.05). ASMI: Appendicular skeletal muscle index; FFMI: Fat-free mass index; BMI: Body mass index; eGFR: estimated glomerular filtration rate; BUN: Blood urea nitrogen; PTH: Parathyroid hormone; CRP: C-reactive protein; T2DM: Type 2 diabetes mellitus; HbA1c: Glycated haemoglobin A1c

while those not normally distributed were reported as median (interquartile range, IQR). For intergroup comparisons of continuous variables, the Mann-Whitney U test was utilised, and Fisher's exact test was applied to categorical variables. Univariate logistic regression was conducted to examine variables associated with sarcopenia, with results presented as odds ratios (ORs) and 95% confidence intervals (95% CIs). The Spearman correlation coefficient was used to analyse correlations between continuous variables. A p-value of less than 0.05 was considered statistically significant. A post-hoc power analysis conducted using G*Power 3.1.9.7 (Faul et al., 2007) indicated that, for the primary comparison, the achieved power with the current sample was 84.2% (α=0.05, two-tailed).[17] For the prevalence estimate of 22.5%, a minimum sample of n=67 was required to achieve a ±10% margin of error at 95% confidence, which the current cohort satisfies. All statistical analyses were conducted using SPSS version 31.0 (IBM Corp., Armonk, NY, USA).

RESULTS

General Characteristics

The mean age of the 71 included patients was 71.8±4.4 years, and 45.1% (n=32) were male. The median eGFR was 42.1 mL/min/1.73 m² (IQR 32.3–50.4). The most common CKD aetiology was T2DM (49.3%), followed by hypertension (38.0%). The median number of comorbidities was 3 (IQR 2–4). The median number of medications was 6 (IQR 4–7). Table 1 presented the demographic, biochemical, and clinical characteristics.

Table 2. Univariate logistic regression analysis for factors associated with sarcopenia

Variable	OR	95% CI	p
BMI (kg/m²)	0.801	0.677–0.947	0.009
T2DM (etiology)	2.842	0.870–9.283	0.084
Male sex	0.321	0.092–1.121	0.075
eGFR <30 vs ≥30 mL/min/1.73 m²	2.700	0.795–9.166	0.111
Haemoglobin (g/dL)	0.818	0.597–1.122	0.213
Proteinuria (mg/day)	0.999	0.997–1.001	0.236
eGFR (mL/min/1.73 m²)	0.979	0.942–1.017	0.271
CRP (mg/L)	1.048	0.970–1.132	0.237
Albumin (g/L)	1.014	0.928–1.108	0.751
Age (years)	0.963	0.845–1.097	0.568

P<0.05 indicates statistical significance. OR: Odds ratio; CI: Confidence interval; BMI: Body mass index; T2DM: Type 2 diabetes mellitus; eGFR: Estimated glomerular filtration rate; CRP: C-reactive protein

Sarcopenia Prevalence

The overall prevalence of sarcopenia was 22.5% (n=16). According to EWGSOP2 staging, 14.1% (n=10) met the criteria for defined sarcopenia and 8.5% (n=6) for probable sarcopenia. BMI was significantly lower in sarcopenic patients (26.2 vs 31.9 kg/m²; p=0.009). ASMI (5.9 vs 8.0 kg/m²; p<0.001) and FFMI (16.2 vs 19.2 kg/m²; p<0.001) were significantly lower in the sarcopenic group, as was handgrip strength (16.0 vs 28.0 kg; p=0.002). Although haemoglobin was numerically lower in the sarcopenic group, the difference did not reach statistical significance (11.9 vs 12.7 g/dL; p=0.155). Proteinuria levels showed a trend towards lower values in sarcopenic patients (70.0 vs 165.0 mg/day; p=0.055). Significant differences were not observed in eGFR, creatinine, albumin, or CRP values between groups.

Factors Associated with Sarcopenia

On univariate logistic regression, BMI was the only variable significantly associated with sarcopenia (OR=0.801; 95% CI: 0.677–0.947; p=0.009). T2DM aetiology (OR=2.842; 95% CI: 0.870–9.283; p=0.084) and male sex (OR=0.321; 95% CI: 0.092–1.121; p=0.075) showed borderline associations. eGFR, haemoglobin, proteinuria, and age were not significantly associated with sarcopenia in univariate analysis. The binary classification of eGFR as CKD eGFR<30 vs. ≥30 mL/min/1.73 m² also did not reach significance (OR=2.700; 95% CI: 0.795–9.166; p=0.111). Logistic regression results are presented in Table 2.

Correlation Analyses

Spearman correlation analysis revealed a strong positive correlation between haemoglobin and ASMI (r=+0.524; p<0.001) and a moderate positive correlation between haemoglobin and handgrip strength (r=+0.437; p<0.001). A strong positive correlation was also observed between BMI and ASMI (r=+0.498; p<0.001). A significant positive correlation was identified between proteinuria and ASMI (r=+0.276; p=0.033). Correlation findings are summarised in Table 3.

Table 3. Correlation analysis between haemoglobin, BMI, proteinuria and sarcopenia components

Variables	Spearman r	p
Haemoglobin –ASMI	+0.524	<0.001
Haemoglobin – Handgrip strength	+0.437	<0.001
Haemoglobin – FFMI	+0.282	0.032
Haemoglobin – Gait speed	+0.307	0.012
BMI – ASMI	+0.498	<0.001
BMI – Handgrip strength	+0.305	0.011
Proteinuria – ASMI	+0.276	0.033
Proteinuria – eGFR	−0.132	0.278

P<0.05 indicates statistical significance. ASMI: Appendicular skeletal muscle index; FFMI: Fat-free mass index; BMI: Body mass index; eGFR: Estimated glomerular filtration rate

DISCUSSION

In this cross-sectional study, sarcopenia prevalence among geriatric patients with predialysis CKD was found to be 22.5% based on the EWGSOP2 criteria. Low body mass index (BMI) emerged as the sole independent factor significantly linked to sarcopenia. Notably, a significant correlation was found between haemoglobin levels and both muscle mass and muscle strength.

The observed prevalence of 22.5% is consistent with the previously reported range of 5–40% in studies focusing on predialysis CKD patients. For example, Souza et al.[13] reported a prevalence of 28.3% using the EWGSOP criteria in a similar cohort. Furthermore, data compiled by Sabatino et al.[12] demonstrate that prevalence rates vary significantly based on measurement methods and patient characteristics. A comprehensive systematic review and meta-analysis by Duarte et al.[14] estimated the prevalence of sarcopenia in CKD patients to be approximately 21%. In Türkiye, Tufan et al.[18] identified low handgrip strength, potentially indicative of sarcopenia, in 29.7% of patients with eGFR<60 mL/min/1.73 m²—a finding consistent with our results and highlighting the clinical burden of muscle weakness in the Turkish CKD population. The relatively lower prevalence observed compared to some series is primarily due to the outpatient setting and the absence of a significant burden of advanced

comorbidities. When only defined sarcopenia cases are considered, prevalence falls to 14.1%, which is consistent with predialysis cohorts reported in the literature.

The lack of a significant association between eGFR level and sarcopenia prevalence is a recurrent finding in comparable studies.[12,19] Although the proportion of sarcopenic patients was numerically higher in those with eGFR<30 mL/min/1.73 m^2 (37.5%) than in those with eGFR≥30 (18.2%), the difference did not reach statistical significance, likely due to the limited number of Stage 4–5 patients. Importantly, muscle loss emerges from the combined effects of the uraemic environment, chronic inflammation, and nutritional disturbances that begin from the early stages of CKD, regardless of eGFR level; as a result, no sharp distinction in sarcopenia risk forms between CKD stages.[6,19]

The strong positive correlation between haemoglobin and both ASMI and handgrip strength is a clinically important finding. Several mechanisms may underlie this relationship. First, lower haemoglobin levels can impair aerobic muscle metabolism, leading to energy insufficiency and accelerated muscle fatigue. Hypoxia causes decreased oxygen delivery to the muscles, and this is more pronounced in anaemia associated with CKD.[20] Second, systemic inflammation, erythropoietin resistance, and iron deficiency contribute in parallel to both renal anaemia[21] and skeletal muscle loss. Third, emerging evidence suggests that correction of renal anaemia may be associated with improved exercise capacity and muscle function in CKD patients; however, the direct effect of erythropoiesis-stimulating agents on skeletal muscle mass and strength warrants further investigation in prospective studies.[22,23] These observations collectively suggest that optimal management of renal anaemia should be considered an integral component of sarcopenia prevention in CKD.

The aetiology of T2DM was observed at a marginally higher frequency within the sarcopenic group, indicating that diabetes-related mechanisms, such as insulin resistance, hyperglycaemic myotoxicity, and peripheral neuropathy, may collectively expedite muscle loss in this demographic. Numerous studies have documented elevated rates of sarcopenia in diabetic patients with CKD compared to their non-diabetic counterparts, with insulin resistance and impaired insulin-like growth factor 1 (IGF-1) signalling identified as principal contributors.[8,12] The SCOPE study reported a 13% prevalence of sarcopenia among diabetic patients diagnosed with chronic renal failure. Within this study, diabetic participants exhibited a higher incidence of adverse outcomes across all sarcopenia components compared to non-diabetic participants.[24]

The observation that proteinuria tends to be lower in sarcopenic patients initially appears paradoxical. One plausible explanation is that patients with elevated proteinuria in this cohort—particularly those with nephrotic syndrome or glomerulonephritis—may be relatively younger individuals with fewer comorbidities and greater capacity to preserve muscle mass. Moreover, concomitant oedema and fluid retention may cause measurement bias in lean mass estimates by BIA. The significant positive correlation between proteinuria and ASMI supports this interpretation. The limited number of patients with high-grade proteinuria in this cohort may have contributed to this counterintuitive finding, and it should be interpreted with caution until confirmed in larger prospective studies with specific proteinuria subgroups.[8]

Body Mass Index (BMI) emerged as the sole independent predictor of sarcopenia. An increased incidence of sarcopenia was observed in patients with a low BMI, which can be attributed to the effects of mechanical loading and adipokine-mediated anabolic signalling. The study underscores the importance of protein-energy malnutrition, frequently seen in CKD, in the pathogenesis of sarcopenia.[25] Furthermore, the possibility of sarcopenic obesity, defined by diminished muscle mass despite a maintained or elevated BMI, should be considered in this population. This is due to the fact that BMI alone may not adequately represent the risk of sarcopenia in overweight or obese CKD patients.

This study has several limitations. The cross-sectional design restricts the ability to infer causal relationships. The modest sample size of patients diminishes the statistical power to assess the association between eGFR and sarcopenia, as well as to separately analyse the three EWGSOP2 sarcopenia stages (probable, confirmed, severe). Potential confounding factors, such as dietary protein intake and physical activity levels, were not systematically recorded. The single-centre design limits the generalisability of the findings. Lastly, BIA-based muscle mass assessment is influenced by hydration status, presenting a methodological limitation, particularly in patients with significant proteinuria or oedema.

CONCLUSION

This study determined that the prevalence of sarcopenia among geriatric patients with non-dialysis-dependent CKD was 22.5% per EWGSOP2 criteria. Although eGFR stage, haemoglobin level, T2DM aetiology, and proteinuria were not independently associated with sarcopenia in univariate analysis, a strong correlation was observed between haemoglobin and both muscle mass and muscle strength, underscoring

the role of renal anaemia in the pathophysiology of sarcopenia. Low BMI emerged as a significant independent risk indicator, highlighting the necessity of early body composition assessment in CKD patients. These findings advocate for routine sarcopenia screening during CKD follow-up. Prospective multicentre studies—especially those examining the impact of anaemia management and diabetes control on muscle mass—are recommended.

Disclosures

Ethics Committee Approval: The study was approved by the University of Health Sciences Kanuni Sultan Suleyman Training and Research Hospital Ethics Committee (No: 2023.07.89, Date: 13/07/2023).

Informed Consent: Written informed consents were obtained from patients who participated in this study.

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

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

Use of AI for Writing Assistance: The authors declared that artificial intelligence was not used in the study.

Author Contributions: Concept – T.E.G., O.K.Ö., M.O., A.Ş., M.A.; Design – T.E.G., O.K.Ö., M.O,. A.S., M.A.; Supervision – T.E.G., M.O., M.A.; Funding – T.E.G., O.K.Ö., M.O., A.S., M.A.; Materials – T.E.G., O.K.Ö., M.O., A.S., M.A.; Data collection and/or processing – T.E.G., O.K.Ö., M.O., A.S., M.A.; Data analysis and/or interpretation – T.E.G., O.K.Ö.; Literature search – T.E.G., O.K.Ö., M.O.; Writing – T.E.G., O.K.Ö., M.O.; Critical review – T.E.G., O.K.Ö., M.O., A.S., M.A.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Cruz-Jentoft AJ, Bahat G, Bauer J, Boirie Y, Bruyère O, Cederholm T, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing. 2019 1;48:16–31. [CrossRef]
2. Bahat G, Tufan A, Tufan F, Kilic C, Akpinar TS, Kose M, et al. Cut-off points to identify sarcopenia according to European Working Group on Sarcopenia in Older People (EWGSOP) definition. Clin Nutr 2016;35:1557–63. [CrossRef]
3. Afkarian M, Sachs MC, Kestenbaum B, Hirsch IB, Tuttle KR, Himmelfarb J, et al. Kidney disease and increased mortality risk in type 2 diabetes. J Am Soc Nephrol 2013;24:302–8. [CrossRef]
4. Zhang QL, Rothenbacher D. Prevalence of chronic kidney disease in population-based studies: systematic review. BMC Public Health 2008;8:117. [CrossRef]
5. Süleymanlar G, Utaş C, Arinsoy T, Ateş K, Altun B, Altiparmak MR, et al. A population-based survey of Chronic REnal Disease In Turkey--the CREDIT study. Nephrol Dial Transplant 2011;26:1862–71. [CrossRef]
6. Wang XH, Mitch WE. Mechanisms of muscle wasting in chronic kidney disease. Nat Rev Nephrol 2014;10:504–16. [CrossRef]
7. Fahal IH. Uraemic sarcopenia: aetiology and implications. Nephrol Dial Transplant 2014;29:1655–65. [CrossRef]
8. An JN, Kim JK, Lee HS, Kim SG, Kim HJ, Song YR. Late stage 3 chronic kidney disease is an independent risk factor for sarcopenia, but not proteinuria. Sci Rep 2021;11:18472. [CrossRef]
9. Pereira RA, Cordeiro AC, Avesani CM, Carrero JJ, Lindholm B, Amparo FC, et al. Sarcopenia in chronic kidney disease on conservative therapy: prevalence and association with mortality. Nephrol Dial Transplant 2015;30:1718–25. [CrossRef]
10. Vettoretti S, Caldiroli L, Armelloni S, Ferrari C, Cesari M, Messa P. Sarcopenia is associated with malnutrition but not with systemic inflammation in older persons with advanced CKD. Nutrients 2019;11:1378. [CrossRef]
11. Hotta C, Hiraki K, Wakamiya A, Otobe Y, Watanabe S, Izawa KP, et al. Relation of physical function and physical activity to sarcopenia in hemodialysis patients: A preliminary study. Int J Cardiol 2015;191:198–200. [CrossRef]
12. Sabatino A, Cuppari L, Stenvinkel P, Lindholm B, Avesani CM. Sarcopenia in chronic kidney disease: what have we learned so far? J Nephrol 2021;34:1347–72. [CrossRef]
13. Souza VA, Oliveira D, Barbosa SR, Corrêa JODA, Colugnati FAB, Mansur HN, et al. Sarcopenia in patients with chronic kidney disease not yet on dialysis: Analysis of the prevalence and associated factors. PLoS One 2017;12. [CrossRef]
14. Duarte MP, Almeida LS, Neri SGR, Oliveira JS, Wilkinson TJ, Ribeiro HS, et al. Prevalence of sarcopenia in patients with chronic kidney disease: a global systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2024;15:501–12. [CrossRef]
15. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med 2009;150:604–12. [CrossRef]
16. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2012 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int Suppl 2013;3:1–150.
17. Faul F, Erdfelder E, Lang AG, Buchner A. G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 2007;39:175–91. [CrossRef]
18. Tufan A, Tufan F, Akpinar TS, Ilhan B, Bahat G, Karan MA. Low glomerular filtration rate as an associated risk factor for sarcopenic muscle strength: is creatinine or cystatin C-based estimation more relevant? Aging Male 2017;20:110–4. [CrossRef]
19. Moorthi RN, Avin KG. Clinical relevance of sarcopenia in chronic kidney disease. Curr Opin Nephrol Hypertens 2017;26:219–28. [CrossRef]
20. Liu H, Li Y, Xiong J. The role of hypoxia-inducible factor-1 alpha in renal disease. Molecules 2022;27:7318. [CrossRef]
21. Jin K, Li X, Ma Y, Yang D, Tan X, Sun Q, et al. Risk factors associated with sarcopenia in patients with chronic kidney disease: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle 2026;17. [CrossRef]
22. Clyne N, Jogestrand T. Effect of erythropoietin treatment on physical exercise capacity and on renal function in predialytic uremic patients. Nephron 1992;60:390–6. [CrossRef]
23. Johansen KL, Finkelstein FO, Revicki DA, Gitlin M, Evans C, Mayne TJ. Systematic review and meta-analysis of exercise tolerance and physical functioning in dialysis patients treated with erythropoiesis-stimulating agents. Am J Kidney Dis 2010;55:535–48. [CrossRef]
24. Formiga F, Moreno-González R, Corsonello A, Carlsson A, Ärnlöv J, Mattace-Raso F, et al. Diabetes, sarcopenia and chronic kidney disease; the Screening for CKD among Older People across Europe (SCOPE) study. BMC Geriatr 2022;22:254. [CrossRef]
25. Prado CM, Wells JC, Smith SR, Stephan BC, Siervo M. Sarcopenic obesity: A Critical appraisal of the current evidence. Clin Nutr 2012;31:583–601. [CrossRef]

DOI: 10.14744/cm.2026.15046
Compreh Med 2026;18(Suppl. 1):53-66

RESEARCH ARTICLE

Many Videos Demonstrate the Operation, Few Teach the Operation: Educational Completeness and Workflow Representation in Oncologic Laparoscopic Adrenalectomy Videos on YouTube

Adem Özcan[1], Gizem Güneş[1], Mahmut Ozan Aydın[2], İhsan Barış Müldür[3]

[1]Department of Surgical Oncology, University of Health Sciences, Ankara Bilkent City Hospital, Ankara, Türkiye
[2]Department of General Surgery, University of Health Sciences, Kanuni Sultan Suleyman Research and Training Hospital, İstanbul, Türkiye
[3]Department of Urology, University of Health Sciences, Ankara Bilkent City Hospital, Ankara, Türkiye

ABSTRACT

Objective: Video-based learning is increasingly used in minimally invasive oncologic surgery; however, whether YouTube videos of oncologic laparoscopic adrenalectomy convey a complete operative workflow remains unclear. This study assessed educational quality, workflow presentation, and educational completeness in such videos.

Materials and Methods: A systematic YouTube search was performed on February 1, 2026, using "laparoscopic adrenalectomy." Of 120 videos screened, 36 met predefined oncologic eligibility criteria. Educational quality was evaluated using DISCERN, JAMA, GQS, and LAP-VEGaS. Educational completeness was assessed using the Adrenalectomy Educational Completeness Score (AECS; 0–9), a procedure-specific score evaluating key workflow components. Associations between scores, video format, narration, and video characteristics were analyzed.

Results: Mean DISCERN, JAMA, GQS, and LAP-VEGaS scores were 55.3±10.5, 1.9±0.9, 3.8±1.0, and 7.1±1.1, respectively. Adrenal vein ligation (86.1%) and dissection (75.0%) were most frequently shown, whereas pathology reporting (33.3%) and discussion of complications or technical challenges (30.6%) were less common. The mean AECS was 4.6±2.6 and correlated significantly with DISCERN, GQS, JAMA, and LAP-VEGaS. Workflow-oriented videos had higher AECS than highlight and partial-operation videos (7.75±1.89 vs. 4.25±2.26 and 4.25±2.49; p=0.044). Narrated videos also showed higher educational completeness (p<0.001). Popularity metrics were not associated with educational quality.

Conclusion: YouTube videos of oncologic laparoscopic adrenalectomy vary in educational completeness. Technical steps are commonly demonstrated, but workflow-defining and oncologic safety-related elements are often underrepresented. Workflow-oriented structure and narration may improve educational value.

Keywords: Educational completeness, laparoscopic adrenalectomy, oncologic adrenal surgery, procedural reproducibility, surgical education, video-based learning, workflow completeness, YouTube

How to cite this article: Özcan A, Güneş G, Aydın MO, Müldür İB. Many Videos Demonstrate the Operation, Few Teach the Operation: Educational Completeness and Workflow Representation in Oncologic Laparoscopic Adrenalectomy Videos on YouTube. Compreh Med 2026;18(Suppl. 1):53-66.

INTRODUCTION

The increasing adoption of minimally invasive techniques in adrenal surgery has expanded the role of laparoscopic adrenalectomy in the management of both benign and selected malignant adrenal lesions.[1–4] Although laparoscopic adrenalectomy offers well-established perioperative advantages, its application in oncologic settings remains particularly demanding because oncologic adequacy depends not only on successful tumor removal but also on meticulous adherence to fundamental surgical principles.[2–5] Avoidance of capsular disruption, precise vascular control, appropriate specimen handling, and maintenance of oncologic integrity throughout the procedure are critical determinants of surgical quality and long-term outcomes.[2–5] Consequently,

Address for Correspondence: Adem Özcan, Department of Surgical Oncology, University of Health Sciences, Ankara Bilkent City Hospital, Ankara, Türkiye
E-mail: dr.ademozcan@gmail.com **ORCID ID:** 0000-0003-2432-7334

Received date: 04.07.2026
Revised date: 02.08.2026
Accepted date: 17.08.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

oncologic laparoscopic adrenalectomy should be viewed not merely as a technical operation but as a complex sequence of workflow-dependent decisions requiring procedural reproducibility, operative judgment, and structured surgical training. In this context, the educational value of a surgical video cannot be judged solely by whether key technical maneuvers are demonstrated; rather, it should be assessed according to its ability to convey the operative workflow, decision-making process, and safety-critical principles that underpin oncologically sound surgery.

The role of minimally invasive surgery in adrenal malignancy remains a subject of ongoing debate, making the quality of surgical education particularly relevant in this field.[2–5] Although laparoscopic adrenalectomy has been shown to be feasible in carefully selected patients with localized adrenocortical carcinoma, metastatic adrenal disease, and other potentially malignant lesions, concerns persist regarding capsular disruption, margin status, tumor spillage, local recurrence, and long-term oncologic outcomes.[2–5] These concerns highlight a fundamental principle: oncologic adrenalectomy is not merely a sequence of technical maneuvers but a procedure in which operative planning, trocar strategy, exposure, dissection planes, vascular control, specimen retrieval, and intraoperative decision-making collectively determine oncologic safety and procedural success.[1–5] Consequently, educational resources intended to support training in oncologic laparoscopic adrenalectomy should extend beyond the demonstration of isolated technical steps and provide a comprehensive representation of the operative workflow. Procedural reproducibility in oncologic surgery depends not only on technical execution but also on understanding how individual operative steps are integrated into a coherent workflow. Whether publicly available surgical videos adequately convey these workflow-defining educational elements remains largely unknown.

Video-based learning has become a cornerstone of contemporary surgical education, particularly in minimally invasive and oncologic surgery, where procedural understanding increasingly relies on the visual demonstration of complex operative workflows.[6,7] Surgical trainees and practicing surgeons routinely use operative videos to review anatomy, understand procedural sequencing, anticipate technical difficulties, and refine operative decision-making before undertaking challenging procedures.[6–8] Among available platforms, YouTube has emerged as one of the most frequently accessed sources of surgical videos because of its unrestricted accessibility, extensive content volume, and global reach.[7–9] However, unlike peer-reviewed video libraries and society-endorsed educational repositories, YouTube content is not required to adhere to standardized reporting criteria, demonstrate the complete operative workflow, disclose outcomes, or undergo expert educational review before publication.[8–10] Consequently, substantial variability exists in the quality, reliability, and educational completeness of publicly available surgical videos.[8–11] More importantly, videos that successfully demonstrate isolated technical maneuvers may still fail to convey the broader workflow, operative context, and decision-making processes required for safe and reproducible oncologic surgery. As procedural reproducibility increasingly becomes a central objective of surgical training, educational value should be assessed not only through global quality metrics but also through the completeness with which a video communicates the operative workflow.

Several instruments have been developed to evaluate the quality and reliability of online surgical videos. DISCERN, the Journal of the American Medical Association (JAMA) benchmark criteria, and the Global Quality Score (GQS) are among the most frequently used tools for assessing the transparency, reliability, and overall educational value of medical content.[12,13] More specifically, the LAParoscopic surgery Video Educational GuidelineS (LAP-VEGaS) assessment tool was developed and validated to provide a structured framework for evaluating the educational quality of laparoscopic surgical videos.[14] Although these instruments have substantially improved the objective assessment of surgical video content, they predominantly generate global quality metrics. In the context of oncologic laparoscopic adrenalectomy, however, educational adequacy extends beyond overall quality scores. A video may achieve acceptable ratings while still failing to demonstrate critical workflow-defining elements such as patient positioning, trocar strategy, anatomical orientation, specimen retrieval, pathology reporting, or the management of intraoperative challenges. Consequently, global quality assessment alone may not adequately reflect whether a video teaches the procedure as a reproducible and oncologically sound operation. As procedural reproducibility increasingly becomes a central objective of modern surgical training, evaluating educational completeness at the procedure-step level may provide a more meaningful assessment of the true educational value of surgical videos.

This distinction is particularly important in oncologic adrenal surgery. A video may achieve acceptable global quality scores while still failing to demonstrate critical workflow-defining components, including patient positioning, trocar placement, anatomical orientation, operative indication, specimen retrieval, pathology reporting, and the management of intraoperative challenges. Conversely, a visually

appealing highlight video may successfully showcase adrenal vein dissection and ligation yet provide little insight into operative planning, procedural sequencing, technical decision-making, or strategies required to maintain oncologic integrity. For malignant adrenal lesions, such omissions extend beyond educational limitations and may compromise the transmission of fundamental oncologic surgical principles. Importantly, the educational objective of an oncologic surgical video is not merely to demonstrate individual technical steps but to communicate how those steps are integrated into a safe, reproducible, and oncologically sound operative workflow. Consequently, assessment of educational value should incorporate not only overall quality metrics but also the completeness with which a video represents the operative workflow of oncologic adrenalectomy.

Although previous studies have evaluated online surgical videos using global quality assessment tools and engagement metrics, relatively little attention has been paid to procedure-step completeness and operative workflow representation.[8–11,14–16] While these investigations have consistently demonstrated substantial heterogeneity in the quality of publicly available surgical content, they provide limited insight into whether videos adequately convey the workflow-defining elements required for reproducible and oncologically sound laparoscopic adrenalectomy. Furthermore, the educational implications of video format—highlight clips, partial procedures, and workflow-oriented videos—remain poorly characterized. This knowledge gap is clinically important because the distinction between a highlight video and a comprehensive workflow video may determine whether the viewer acquires an isolated technical maneuver or a meaningful understanding of the operation as a complete oncologic procedure. Consequently, educational completeness and workflow representation may provide a more clinically relevant assessment of surgical video quality than global quality metrics alone.

Therefore, the present study aimed to evaluate YouTube videos demonstrating oncologic laparoscopic adrenalectomy using established educational quality assessment instruments while simultaneously assessing operative workflow representation. In addition to DISCERN, JAMA, GQS, and LAP-VEGaS scores, we evaluated the representation of workflow-defining educational components, including patient positioning, trocar placement, anatomical explanation, adrenal vein dissection and ligation, specimen extraction, operative indication, pathology reporting, and discussion of complications or technical challenges. Furthermore, we assessed educational completeness using a novel procedure-specific metric, the Adrenalectomy Educational Completeness Score (AECS), and investigated its relationship with video format and conventional educational quality measures. We hypothesized that although technical dissection-related steps would be frequently represented, workflow-defining and oncologic safety-related educational elements would remain substantially underrepresented, particularly in highlight and partial-operation videos.

MATERIALS and METHODS

Study Design and Ethical Considerations

This study was designed as a retrospective cross-sectional observational analysis to evaluate the educational quality, workflow representation, and educational completeness of YouTube videos demonstrating oncologic laparoscopic adrenalectomy. In addition to conventional video quality assessment instruments, the study specifically focused on the representation of workflow-defining educational components considered essential for procedural reproducibility and oncologic surgical practice.

All videos analyzed in this study were obtained from YouTube (https://www.youtube.com), a publicly accessible online platform. The study involved neither direct human participation nor access to patient-level data, personal identifiers, or protected health information. Because the investigation consisted exclusively of the evaluation of publicly available educational content, ethics committee approval and informed consent were not required.

Search Strategy and Video Selection

A systematic search was performed on YouTube on February 1, 2026, using the search term "laparoscopic adrenalectomy." Search results were screened according to YouTube's default relevance algorithm to reflect the typical experience of users seeking educational surgical content.

Initially, 120 videos were identified and reviewed for eligibility. All videos were screened according to predefined inclusion and exclusion criteria. Duplicate videos, patient-oriented educational videos, non-operative content, videos unrelated to laparoscopic adrenalectomy, videos demonstrating clearly benign adrenal procedures without oncologic relevance, and videos with insufficient visual quality were excluded.

Of the 120 initially identified videos, 84 were excluded, including duplicate videos (n=13), patient-oriented educational videos (n=27), benign or non-oncologic adrenal procedures (n=28), videos with insufficient visual quality for detailed assessment (n=12), and non-operative videos (n=4). Following the screening process, 36 videos demonstrating malignant, potentially malignant, metastatic, or otherwise oncologically

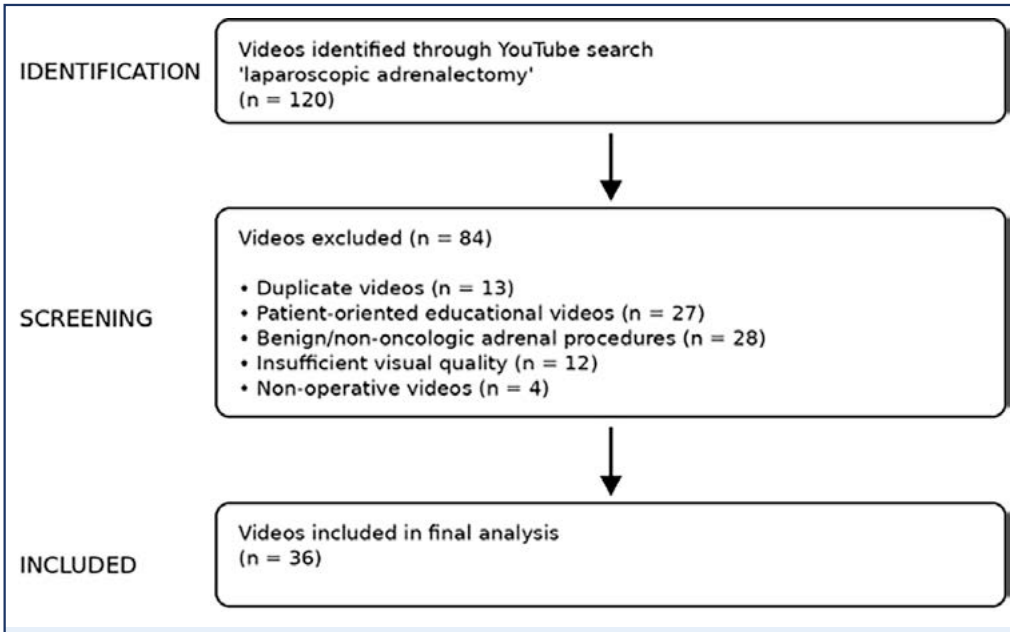


Figure 1. Flow diagram of video selection

managed adrenal lesions met all eligibility criteria and were included in the final analysis (Fig. 1).

Inclusion and Exclusion Criteria

Videos were eligible for inclusion if they:

- Demonstrated laparoscopic adrenalectomy performed for malignant, potentially malignant, metastatic, or otherwise oncologically managed adrenal lesions;
- Were intended for surgical education;
- Contained operative footage;
- Provided sufficient visual quality for detailed evaluation;
- Contained sufficient information within the title, description, narration, or video content to establish an oncologic surgical context.

Videos were excluded if they:

- Were duplicates;
- Did not contain operative content;
- Were designed exclusively for patient education or public information;
- Consisted solely of animations, lectures, interviews, or non-operative presentations;
- Demonstrated clearly benign adrenal procedures without oncologic relevance;
- Lacked sufficient information to establish an oncologic context;
- Had insufficient visual quality for objective assessment.

Oncologic Case Selection

The primary objective of this study was to evaluate educational content relevant to oncologic adrenal surgery. Therefore, only videos demonstrating malignant adrenal tumors, metastatic adrenal lesions, pheochromocytomas managed within an oncologic surgical context, or adrenal masses explicitly presented within an oncologic surgical context were included.

Because pathological confirmation was not independently available, classification relied on information provided by the uploaders. Accordingly, the term "oncologic laparoscopic adrenalectomy" in the present study refers to procedures presented as malignant, potentially malignant, metastatic, or otherwise managed according to oncologic surgical principles.

Data Collection

For each included video, the following characteristics were recorded:

- Video duration (seconds);
- Number of views;
- Number of likes;
- Number of comments;
- Upload date;
- Video age (days);
- Uploader type;
- Laterality (right or left adrenalectomy);
- Presence of subtitles;

- Presence of voice narration;
- Video format.

Video age was calculated as the number of days between the upload date and the date of data collection.

Uploader type was categorized as either an individual surgeon (coded as 1) or an institutional source (coded as 2). Institutional sources included universities, hospitals, professional organizations, academic centers, and officially affiliated educational channels.

Video Format Classification

Videos were categorized according to the extent of operative workflow representation:

- Highlight video: videos demonstrating selected operative moments or key technical maneuvers without comprehensive representation of the overall procedure.
- Partial-operation video: videos demonstrating substantial portions of the operation but omitting one or more major workflow-defining stages.
- Workflow-oriented video: videos presenting the procedure comprehensively, including operative setup, key technical steps, and procedural completion.

Video format classification was independently performed by the reviewers based on the extent to which the complete operative workflow was represented. Disagreements were resolved through consensus discussion.

Procedure-Step Assessment

The presence or absence of workflow-defining educational components was assessed for each video. Components were selected based on their relevance to operative orientation, technical execution, procedural reproducibility, specimen management, and oncologic surgical practice.

The following variables were evaluated:

- Patient positioning;
- Trocar placement;
- Anatomical explanation;
- Adrenal vein dissection;
- Adrenal vein ligation;
- Specimen extraction;
- Operative indication;
- Pathology reporting;
- Discussion of complications or technical challenges.

Additionally, the presence of voice narration and subtitles was recorded.

Each variable was assessed dichotomously and recorded as absent (0) or present (1).

Adrenalectomy Educational Completeness Score (AECS)

To evaluate educational completeness beyond conventional quality metrics, an Adrenalectomy Educational Completeness Score (AECS) was developed specifically for this study.

The AECS consisted of nine workflow-defining components:

1. Patient positioning;
2. Trocar placement;
3. Anatomical explanation;
4. Adrenal vein dissection;
5. Adrenal vein ligation;
6. Specimen extraction;
7. Operative indication;
8. Pathology reporting;
9. Discussion of complications or technical challenges.

The AECS components were selected a priori by the study investigators based on their relevance to procedural reproducibility, operative orientation, technical execution, specimen management, and oncologic safety in laparoscopic adrenalectomy. Component selection was informed by the educational domains emphasized in LAP-VEGaS and by workflow-defining elements considered essential for understanding and reproducing an oncologic minimally invasive adrenal procedure. The aim of the AECS was not to replace established quality assessment instruments but to provide a procedure-specific measure of educational completeness.

Each component contributed one point if present. Consequently, total AECS values ranged from 0 to 9, with higher scores indicating greater educational completeness and a more comprehensive representation of the operative workflow.

Educational Quality Assessment

Educational quality was independently evaluated by three reviewers experienced in minimally invasive and oncologic surgery. The following validated assessment tools were used: the DISCERN score, the Journal of the American Medical Association (JAMA) benchmark criteria, the Global Quality Score (GQS), and the LAParoscopic surgery Video Educational GuidelineS (LAP-VEGaS). All assessments were performed independently by the three reviewers. For each video, the final DISCERN, JAMA, GQS, and LAP-VEGaS scores were calculated as the mean values of the three independent evaluations.

Interobserver Reliability

Interobserver agreement among reviewers for LAP-VEGaS scoring was assessed using the intraclass correlation coefficient (ICC), as LAP-VEGaS required the greatest degree of subjective reviewer judgment among the assessment instruments employed. ICC values greater than 0.75 were considered indicative of good agreement, whereas values greater than 0.90 were considered indicative of excellent agreement.

Use of Artificial Intelligence–Assisted Technology

Artificial intelligence–assisted technology was used only for editorial and language-related support during manuscript preparation. Specifically, ChatGPT (OpenAI, GPT-5.5 Pro) was used to improve language clarity, assist with Turkish–English translation, shorten the abstract according to journal word limits, refine the manuscript title, and draft submission-related text. Representative prompts included requests to translate the abstract into Turkish, limit the Turkish and English abstracts to 250 words, refine the title, and prepare a cover letter emphasizing the contribution of the study to the literature. The tool was not used to identify or select videos, determine eligibility, extract study data, assign DISCERN, JAMA, GQS, LAP-VEGaS, or AECS scores, perform statistical analyses, generate tables or figures, interpret the results, or draw scientific conclusions. All AI-assisted outputs were reviewed, edited, and approved by the authors. The authors take full responsibility for the accuracy, integrity, and originality of the final manuscript.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 22.0 (IBM Corp., Armonk, NY, USA).

The distribution of continuous variables was assessed using the Shapiro–Wilk test. Continuous variables were presented as mean±standard deviation or median and interquartile range (IQR), as appropriate.

Comparisons between two groups were performed using the Mann–Whitney U test. Comparisons among video format categories (highlight, partial-operation, and workflow-oriented videos) were performed using the Kruskal–Wallis test.

Correlations between educational quality metrics and educational completeness measures, including AECS, DISCERN, JAMA, GQS, LAP-VEGaS, video duration, views, likes, and comments, were evaluated using Spearman's rank correlation coefficient.

Categorical variables were compared using the chi-square test or Fisher's exact test, as appropriate.

All statistical tests were two-sided, and a p-value<0.05 was considered statistically significant.

Table 1. General characteristics of included videos (n=36)

Variable	Value
Video duration, seconds	412 (293–511)
Views	233.5 (102.3–956.3)
Likes	4 (1–16.3)
Comments	0 (0–1)
Video age, days	1065.5 (548–1626.8)
Individual surgeon upload, n (%)	24 (66.7)
Institutional upload, n (%)	12 (33.3)
Left adrenalectomy, n (%)	20 (55.6)
Right adrenalectomy, n (%)	16 (44.4)
Highlight video, n (%)	12 (33.3)
Partial-operation video, n (%)	20 (55.6)
Workflow-oriented video, n (%)	4 (11.1)
Voice narration present, n (%)	23 (63.9)
Subtitles present, n (%)	15 (41.7)

Data are presented as median (interquartile range) or number (percentage), as appropriate

RESULTS

Video Selection

A total of 120 videos were initially identified through the YouTube search. After application of the predefined inclusion and exclusion criteria, 84 videos were excluded, including duplicate videos (n=13), patient-oriented educational videos (n=27), benign or non-oncologic adrenal procedures (n=28), videos with insufficient visual quality for detailed assessment (n=12), and non-operative videos (n=4). Consequently, 36 videos demonstrating malignant, potentially malignant, metastatic, or otherwise oncologically managed adrenal lesions met all eligibility criteria and were included in the final analysis (Fig. 1).

General Characteristics of Included Videos

The general characteristics of the included videos are presented in Table 1.

The median video duration was 412 seconds (IQR: 293–511). The median number of views was 233.5 (IQR: 102.3–956.3), while the median numbers of likes and comments were 4 (IQR: 1–16.3) and 0 (IQR: 0–1), respectively. The median video age at the time of analysis was 1065.5 days (IQR: 548–1626.8).

Most videos were uploaded by individual surgeons (n=24, 66.7%), whereas 12 videos (33.3%) originated from institutional sources. Twenty videos (55.6%) demonstrated left adrenalectomy, and 16 (44.4%) demonstrated right adrenalectomy.

Table 2. Conventional educational quality assessment scores of included videos (n=36)

Variable	Mean±SD	Median (IQR)
DISCERN score	55.3±10.5	54.5 (50–63)
JAMA score	1.9±0.9	2 (1–3)
GQS score	3.8±1.0	4 (3–5)
LAP-VEGaS score	7.1±1.1	7.0 (6.3–7.3)

SD: Standard deviation; DISCERN: Quality Criteria for Consumer Health Information; JAMA: Journal of the American Medical Association benchmark criteria; GQS: Global Quality Score; LAP-VEGaS: LAParoscopic surgery Video Educational GuidelineS

With respect to video format, 12 videos (33.3%) were classified as highlight videos, 20 (55.6%) as partial-operation videos, and only 4 videos (11.1%) were classified as workflow-oriented videos. Voice narration was present in 23 videos (63.9%), whereas subtitles were available in 15 videos (41.7%).

Conventional Educational Quality Assessment

Conventional educational quality metrics are summarized in Table 2. The mean DISCERN, JAMA, GQS, and LAP-VEGaS scores were 55.3±10.5, 1.9±0.9, 3.8±1.0, and 7.1±1.1, respectively. Interobserver agreement for LAP-VEGaS assessment was good, with an average-measures intraclass correlation coefficient (ICC) of 0.80, supporting the reliability of the educational quality assessments performed in this study.

Table 3. Representation of workflow-defining educational components in included videos (n=36)

Category	Educational component	n (%)
Technical	Adrenal vein ligation	31 (86.1)
Technical	Adrenal vein dissection	27 (75.0)
Workflow	Anatomical explanation	21 (58.3)
Workflow	Operative indication	21 (58.3)
Workflow	Patient positioning	15 (41.7)
Workflow	Trocar placement	15 (41.7)
Outcome/Context	Specimen extraction	14 (38.9)
Outcome/Context	Pathology reporting	12 (33.3)
Outcome/Context	Discussion of complications or technical challenges	11 (30.6)

Data are presented as number (%)

Representation of Workflow-Defining Educational Components

Marked heterogeneity was observed in the representation of workflow-defining educational components (Fig. 2 and Table 3). Adrenal vein ligation was the most frequently demonstrated educational component, appearing in 31 of 36 videos (86.1%), followed by adrenal vein dissection in 27 videos

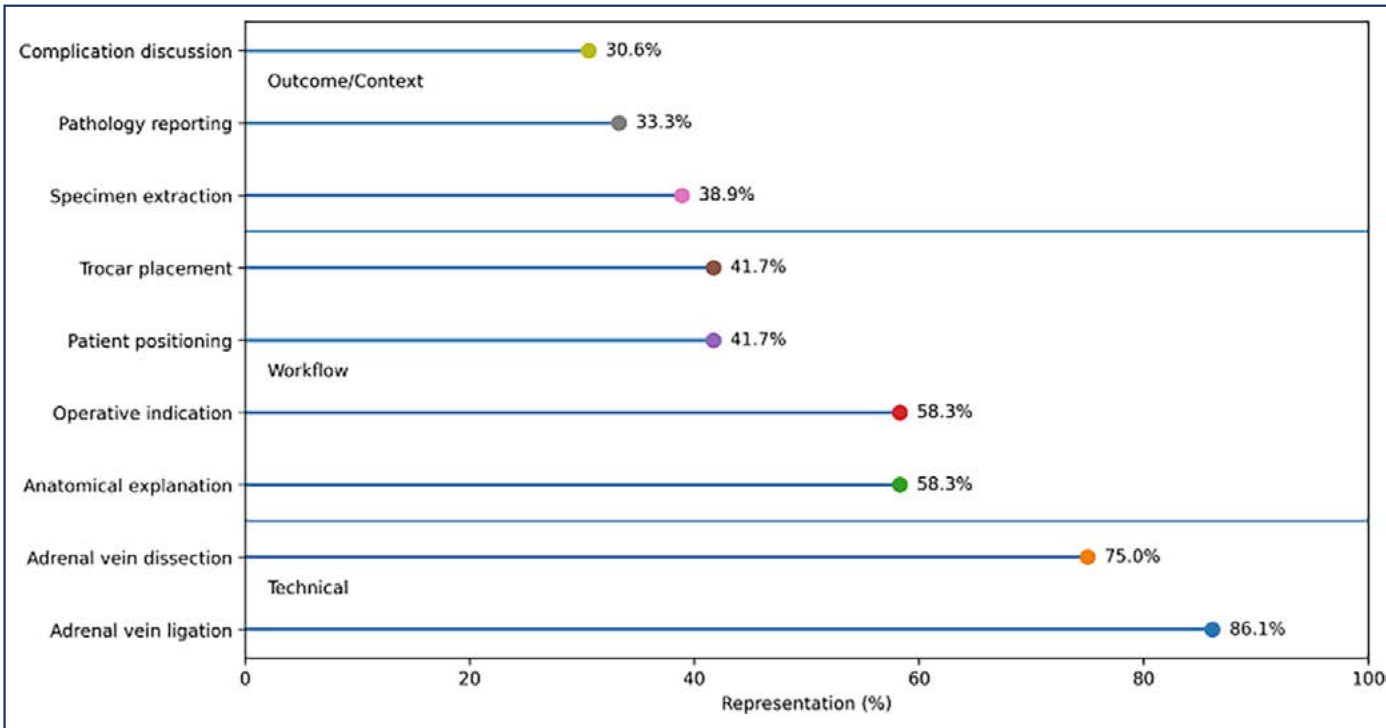


Figure 2. Educational component representation in oncologic laparoscopic adrenalectomy videos

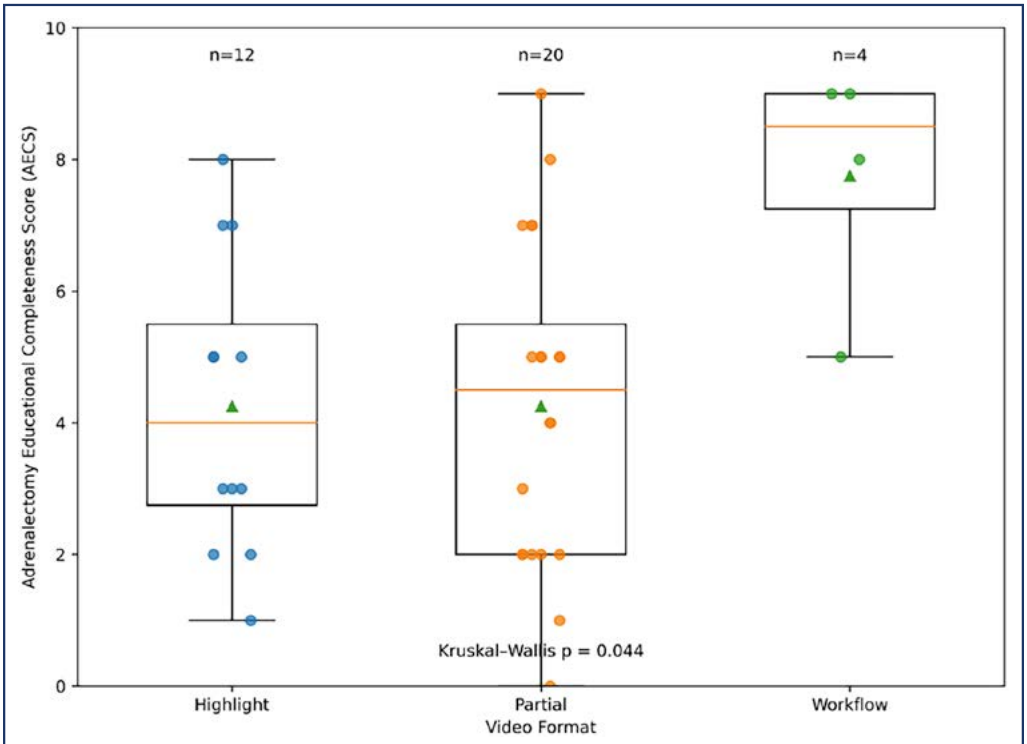


Figure 3. Educational completeness according to video format

(75.0%). Anatomical explanation and operative indication were each demonstrated in 21 videos (58.3%).

In contrast, several educational elements related to operative setup, specimen management, and postoperative context were considerably less frequently represented. Patient positioning and trocar placement were demonstrated in only 15 videos (41.7%), specimen extraction in 14 videos (38.9%), pathology reporting in 12 videos (33.3%), and discussion of complications or technical challenges in 11 videos (30.6%).

A progressive decline in representation was observed from technical dissection-related components to workflow-related and outcome-related educational elements. While technical steps such as adrenal vein dissection and ligation were frequently demonstrated, educational components related to operative preparation, specimen management, pathology, and complication-related decision-making were substantially underrepresented. This pattern suggests that technical execution was more consistently represented than broader workflow-defining educational content.

Adrenalectomy Educational Completeness Score

The mean Adrenalectomy Educational Completeness Score (AECS) was 4.6±2.6, with a median score of 5.0 (IQR: 2–7). AECS values ranged from 0 to 9, indicating substantial variability in the completeness with which operative workflow components were represented across videos.

The AECS demonstrated significant positive correlations with the DISCERN score (ρ=0.682, p<0.001), GQS (ρ=0.558, p<0.001), JAMA score (ρ=0.463, p=0.004), and LAP-VEGaS score (ρ=0.456, p=0.005). Videos with higher AECS values consistently achieved higher scores across established educational quality instruments, supporting the validity of AECS as a procedure-specific measure of educational completeness.

Despite moderate overall quality scores, considerable variability in AECS values was observed, indicating that videos with similar global educational ratings may differ substantially in their representation of workflow-defining educational elements. This finding suggests that educational completeness captures an aspect of video quality that is not fully reflected by conventional assessment instruments alone.

Influence of Video Format on Educational Completeness

Educational completeness differed significantly according to video format (Fig. 3 and Table 4).

Workflow-oriented videos achieved the highest AECS values, with a mean score of 7.75±1.89, compared with 4.25±2.49 for partial-operation videos and 4.25±2.26 for highlight videos

Table 4. Educational quality and educational completeness according to video format

Variable	Highlight (n=12)	Partial- operation (n=20)	Workflow- oriented (n=4)	p
AECS	4.25±2.26	4.25±2.49	7.75±1.89	0.044
DISCERN score	54.42±7.81	53.90±11.87	64.50±7.19	0.153
JAMA score	1.25±0.45	2.15±0.93	3.00±0.00	<0.001
GQS score	3.50±0.90	3.85±1.04	4.50±0.58	0.189
LAP-VEGaS score	6.56±0.67	7.07±0.83	9.08±1.20	0.003

Statistical significance was defined as p<0.05. AECS: Adrenalectomy Educational Completeness Score; DISCERN: Quality Criteria for Consumer Health Information; JAMA: Journal of the American Medical Association benchmark criteria; GQS: Global Quality Score; LAP-VEGaS: LAParoscopic surgery Video Educational GuidelineS. Continuous variables were compared using the Kruskal–Wallis test

(p=0.044). Notably, workflow-oriented videos demonstrated nearly twofold greater educational completeness than both highlight and partial-operation videos.

A similar pattern was observed for LAP-VEGaS scores. Workflow-oriented videos demonstrated the highest educational quality, with a mean LAP-VEGaS score of 9.08±1.20, compared with 7.07±0.83 for partial-operation videos and 6.56±0.67 for highlight videos (p=0.003). JAMA scores also differed significantly according to video format, with workflow-oriented videos achieving the highest values (3.00±0.00), followed by partial-operation videos (2.15±0.93) and highlight videos (1.25±0.45) (p<0.001).

In contrast, no statistically significant differences were observed among video format categories for the DISCERN score, GQS, video duration, or view count. Collectively, these findings suggest that workflow-oriented videos were associated with higher educational completeness and selected quality metrics; however, these exploratory subgroup comparisons should be interpreted cautiously because the workflow-oriented group included only four videos.

Influence of Voice Narration and Subtitles

Videos containing voice narration demonstrated significantly greater educational completeness than videos without narration. The median AECS was 5.0 (IQR: 5–7.5) in narrated videos compared with 2.0 (IQR: 2–3) in non-narrated videos (p<0.001), indicating substantially more comprehensive workflow representation in videos accompanied by verbal explanation.

Narrated videos also achieved significantly higher DISCERN (p=0.010), JAMA (p=0.005), and GQS scores (p=0.044). However, no significant difference in LAP-VEGaS scores was observed according to the presence of voice narration (p=0.219).

In contrast, the presence of subtitles was not significantly associated with AECS, DISCERN, JAMA, GQS, or LAP-VEGaS scores. Overall, videos incorporating voice narration demonstrated greater educational completeness and higher conventional educational quality scores than videos relying solely on visual presentation.

Uploader Type, Laterality, and Popularity Metrics

No significant differences were observed between videos uploaded by individual surgeons and those uploaded by institutional sources with respect to AECS, DISCERN, JAMA, GQS, or LAP-VEGaS scores. Similarly, no significant differences were identified between right- and left-sided adrenalectomy videos regarding educational completeness or educational quality metrics.

View count and comment count were not significantly correlated with either AECS or LAP-VEGaS scores. Like count demonstrated weak positive correlations with AECS (ρ=0.363, p=0.030) and the LAP-VEGaS score (ρ=0.336, p=0.045).

DISCUSSION

The principal finding of this study is that publicly available YouTube videos of oncologic laparoscopic adrenalectomy preferentially emphasize technical execution while frequently underrepresenting workflow-defining educational elements. Although adrenal vein dissection and ligation were demonstrated in the majority of videos, critical components such as patient positioning, trocar placement, specimen extraction, pathology reporting, and discussion of complications or technical challenges were substantially less common. These findings suggest that many videos successfully demonstrate how an adrenalectomy is technically performed, yet comparatively few teach how the procedure is planned, contextualized, and completed as an oncologically sound operation. Consequently, technical exposure and educational completeness should not be considered synonymous, particularly in procedures where procedural reproducibility and oncologic safety depend on a sequence of interconnected operative decisions.

This observation is particularly relevant in adrenal oncology. The role of minimally invasive surgery in malignant adrenal disease remains controversial, particularly in adrenocortical carcinoma and other potentially malignant adrenal lesions.[2–5] Concerns regarding capsular disruption, tumor spillage, margin status, and long-term oncologic outcomes continue to influence surgical decision-making.[2–5] Contemporary adrenal oncology guidelines and outcome studies consistently emphasize that procedural quality in adrenal surgery extends beyond successful tumor removal and requires meticulous adherence to oncologic principles throughout the operative process.[1–5] Consequently, educational resources intended to support oncologic adrenal surgery should ideally communicate not only technical maneuvers but also the sequence of workflow-dependent decisions that contribute to oncologic safety. In this context, the underrepresentation of patient positioning, trocar placement, specimen extraction, pathology reporting, and complication management identified in the present study represents a meaningful educational deficiency rather than a simple limitation of video editing. Indeed, these components are directly related to procedural reproducibility and may influence how surgeons translate operative knowledge into safe oncologic practice. As a result, incomplete workflow representation may affect not only educational quality but also the transmission of fundamental oncologic surgical principles.

A distinguishing feature of the present study is the evaluation of educational completeness at the procedure-step level. Previous investigations examining surgical videos on YouTube have predominantly focused on global quality assessment instruments such as DISCERN, JAMA, GQS, and LAP-VEGaS.[8–16] While these tools provide valuable information regarding transparency, reliability, and overall educational quality, they do not necessarily determine whether a procedure is presented as a complete and reproducible workflow. Similar limitations have been highlighted in recent studies of video-based surgical education, which emphasize that procedural understanding depends not only on technical visualization but also on the representation of operative sequencing, contextual decision-making, and workflow continuity.[17–22]

To address this limitation, we developed the Adrenalectomy Educational Completeness Score (AECS), a procedure-specific metric designed to quantify the representation of workflow-defining educational components. The significant correlations observed between AECS and DISCERN, JAMA, GQS, and LAP-VEGaS support its construct validity and suggest that workflow completeness represents an important educational domain that is only partially captured by existing assessment instruments.[12–14] Notably, the substantial variability observed in AECS values despite relatively similar global quality scores indicates that videos with comparable educational ratings may differ considerably in their ability to communicate a reproducible operative workflow. This finding supports the concept that procedural reproducibility, rather than technical visualization alone, may represent a more meaningful indicator of educational value in oncologic surgical videos.

Importantly, the purpose of AECS was not to evaluate technical complexity or oncologic difficulty but rather to assess whether workflow-defining educational elements were adequately represented in a manner that supports procedural understanding and reproducibility. Consequently, AECS may complement existing educational quality instruments by providing procedure-specific information regarding workflow completeness that cannot be fully captured by global quality metrics alone.

A key exploratory finding of this study was the association between video format and educational value. Workflow-oriented videos achieved significantly higher AECS, LAP-VEGaS, and JAMA scores than both highlight and partial-operation videos. Notably, workflow-oriented videos demonstrated nearly twofold greater educational completeness than highlight videos. Although previous studies have documented substantial heterogeneity in the educational quality of online surgical videos,[9–11,15,16] relatively little attention has been paid to the educational implications of video format itself.

The present findings suggest that educational quality depends not only on what is shown but also on how the procedure is presented. Highlight videos often provide excellent visualization of isolated technical maneuvers; however, they frequently omit procedural context, operative sequencing, and decision-making processes. In contrast, workflow-oriented videos appear to provide a more comprehensive representation of procedural conduct, thereby reducing the cognitive gap between observing an operation and understanding an operation. This observation is consistent with contemporary concepts of competency-based surgical education, which emphasize procedural understanding, cognitive integration, and workflow mastery in addition to technical exposure.[17–22]

Importantly, workflow-oriented videos achieved higher performance across multiple educational domains, including educational completeness, procedural reporting quality, and

transparency. However, because only four videos were classified as workflow-oriented, these findings should be regarded as exploratory and hypothesis-generating rather than definitive. Furthermore, workflow representation appeared to demonstrate stronger associations with educational value than traditional popularity metrics or video duration. Collectively, these findings suggest that workflow completeness may be an important determinant of educational quality in oncologic surgical videos while emphasizing the need for confirmation in larger video datasets.

An additional observation was the progressive decline in representation from technical dissection-related components to workflow-related and outcome-related educational elements. Adrenal vein dissection and ligation were among the most frequently demonstrated procedural steps, whereas pathology reporting and discussion of complications were among the least represented. This pattern likely reflects the way surgical videos are commonly produced. Technical moments are visually appealing, easier to edit, and often perceived as the most educational parts of a procedure. Similar trends have been reported in previous evaluations of online surgical videos, where technical execution was frequently emphasized at the expense of procedural context and decision-making.[9–11,15,16]

However, from a surgical education perspective, pathology reporting, specimen management, and complication discussion may be equally important because they provide insight into operative judgment, oncologic reasoning, and the management of unexpected events. In malignant adrenal disease, these elements are directly related to procedural reproducibility and oncologic safety.[1–5] Indeed, many videos demonstrated the operation, but comparatively few taught the operation. This distinction reflects a fundamental challenge in video-based surgical education: technical visualization alone cannot fully substitute for the cognitive apprenticeship required to understand why a procedure is performed in a particular manner and how operative decisions are made in real time.[17–22] Consequently, educational completeness should be viewed not only as the demonstration of technical steps but also as the communication of operative reasoning, workflow continuity, and context-dependent decision-making.

Voice narration emerged as another important determinant of educational value. Videos containing narration demonstrated significantly greater educational completeness than videos without narration, with a median AECS more than twice as high in narrated videos. Interestingly, subtitles alone were not associated with improved educational quality. This finding suggests that the educational benefit of narration extends beyond simple information delivery. Verbal explanation allows surgeons to communicate operative rationale, anatomical landmarks, procedural sequencing, decision-making processes, and responses to technical difficulties. Similar observations have been reported in the broader surgical education literature, where guided audiovisual instruction consistently outperforms passive observation alone in promoting procedural understanding and skill acquisition.[6,7,17,18,20,22]

The present findings reinforce this concept and suggest that narration may serve as a critical bridge between technical visualization and meaningful educational engagement. Notably, the absence of a similar association for subtitles implies that educational value may depend less on the availability of supplementary information and more on the active communication of operative reasoning. Consequently, voice narration should be considered a core component of future workflow-oriented educational videos in oncologic surgery.

The lack of meaningful associations between educational quality and popularity metrics is also noteworthy. Neither view count nor comment count demonstrated significant correlations with AECS or LAP-VEGaS scores, while like count showed only weak positive associations. These findings are consistent with previous studies demonstrating that online engagement metrics are poor surrogates for educational value and should not be interpreted as indicators of instructional quality.[8–11,15,16] For trainees, this observation is particularly important because it challenges the common assumption that highly viewed videos are necessarily superior educational resources.

In the present study, educational completeness was more strongly associated with workflow representation and voice narration than with visibility or user engagement metrics. This finding suggests that educational design may be a more important determinant of learning value than platform popularity. Consequently, the most visible videos are not necessarily the most educational, and trainees may benefit more from selecting videos based on instructional structure and workflow completeness rather than audience engagement alone. Collectively, these findings indicate that how educational content is organized and delivered may be more important than how widely it is viewed.

Interestingly, institutional videos did not outperform videos uploaded by individual surgeons, and educational quality

did not differ according to operative laterality. These findings suggest that educational value is influenced less by the source of publication and more by the completeness, structure, and instructional design of the educational content itself. Similar variability in educational quality has been reported across different forms of online surgical educational content, where institutional affiliation was not consistently associated with superior educational performance.[9,18,21,23] From an educational perspective, these findings reinforce the notion that effective surgical teaching depends primarily on how procedural knowledge is organized and communicated rather than on who publishes the material. Consequently, institutional affiliation alone should not be considered a guarantee of educational quality.

The broader implications of these findings extend beyond adrenal surgery. Modern surgical training increasingly relies on video-based learning, particularly in minimally invasive and oncologic procedures, where opportunities for direct operative exposure may be limited and procedural complexity is high.[6,7,17,18,20,22] However, if online videos disproportionately emphasize technical execution while neglecting workflow-defining educational elements, trainees may acquire fragmented technical knowledge without fully understanding the procedural logic that underpins safe surgical practice. This concern may be particularly relevant in oncologic surgery, where procedural reproducibility is intimately linked to patient safety, oncologic adequacy, and long-term outcomes.[1–5] The present findings therefore support a shift from evaluating whether videos merely demonstrate technical steps toward assessing whether they communicate a complete operative workflow. From an educational perspective, workflow completeness may represent a more meaningful indicator of instructional value than technical visualization alone and may better reflect a video's capacity to support procedural understanding and reproducible surgical practice.

The findings of this study also support the development of procedure-specific educational standards for online surgical videos. Existing frameworks such as LAP-VEGaS have significantly advanced the assessment of laparoscopic video quality by promoting transparency and standardized reporting.[14] However, the present findings suggest that educational quality should be evaluated not only through global quality metrics but also through the completeness with which workflow-defining elements are represented. Future educational standards may therefore benefit from greater emphasis on patient positioning, trocar strategy, operative context, pathology reporting, complication management, and procedural decision-making.[14,17,19,20,24] Such an approach may improve both the reproducibility and educational utility of surgical videos intended for oncologic training, particularly as social media platforms are increasingly used for surgical education and research dissemination.[25] In addition, randomized evidence in basic surgical skills training supports the educational value of structured video-based learning, reinforcing the need for surgical videos to be designed as instructional resources rather than passive visual demonstrations.[26]

Moreover, the superior performance of workflow-oriented videos across multiple educational domains in the present study provides empirical support for incorporating workflow completeness into future surgical video reporting frameworks. Accordingly, workflow representation should be considered not merely an educational supplement but a core component of high-quality surgical video design.

This study possesses several important strengths. Unlike many previous evaluations of online surgical videos, the present analysis focused specifically on malignant or potentially malignant adrenal lesions, thereby maintaining a clear oncologic perspective throughout the study. This distinction is important because the educational requirements of oncologic adrenalectomy extend beyond technical execution and include procedural reproducibility, oncologic safety, and operative decision-making. Furthermore, educational quality was assessed using both established assessment instruments and a novel procedure-specific framework. By combining DISCERN, JAMA, GQS, and LAP-VEGaS with the Adrenalectomy Educational Completeness Score, we were able to evaluate not only overall educational quality but also the completeness of workflow representation. Additional strengths include independent evaluation by three reviewers and good interobserver agreement. Together, these methodological features provided a comprehensive assessment of both educational quality and workflow completeness in videos demonstrating oncologic laparoscopic adrenalectomy.

Several limitations should be acknowledged when interpreting the present findings. First, YouTube is a dynamic platform in which video availability, metadata, and ranking algorithms may change over time; consequently, the analyzed dataset represents a snapshot of the content available at the time of the search. Second, although the Adrenalectomy Educational Completeness Score (AECS) demonstrated significant correlations with established educational quality instruments, it was developed specifically for this study and should therefore be considered exploratory un-

til externally validated in independent datasets. Third, despite efforts to restrict inclusion to malignant or potentially malignant adrenal lesions, the accuracy of pathology and oncologic context depended on information provided by uploaders and could not be independently verified. Finally, the relatively small number of workflow-oriented videos (n=4) represents an important limitation of the subgroup analyses. Although non-parametric methods were used, the small size of this subgroup may have reduced statistical power and increased the risk of Type II error, particularly for comparisons in which no statistically significant difference was observed. Therefore, the subgroup findings should be interpreted cautiously and considered exploratory. Nevertheless, the small workflow-oriented subgroup also reflects the scarcity of videos meeting predefined oncologic eligibility criteria and presenting a complete operative workflow. Accordingly, the observed distribution likely represents the current state of publicly available educational content rather than a sampling artifact.

CONCLUSION

YouTube videos demonstrating oncologic laparoscopic adrenalectomy exhibit considerable variability in educational completeness. While technical dissection-related components are frequently represented, workflow-defining educational elements—including patient positioning, trocar placement, specimen management, pathology reporting, and complication discussion—remain substantially underrepresented. Workflow-oriented and narrated videos were associated with higher educational completeness, suggesting that educational quality depends not only on technical visualization but also on the completeness with which the operative workflow is communicated.

The findings of this study indicate that procedural reproducibility, rather than technical exposure alone, may represent a meaningful indicator of educational value in oncologic surgical videos. Future educational content should prioritize comprehensive workflow representation, structured narration, and transparent reporting of operative context to better support training in oncologic minimally invasive surgery. In addition, YouTube-based surgical educational content may benefit from a predefined educational format in which the intended video type—highlight, partial-operation, or complete workflow—is specified during recording or upload. Ultimately, workflow completeness should be regarded as a core component of high-quality surgical video design and a key determinant of educational value in oncologic surgical education.

Disclosures

Ethics Committee Approval: This study analyzed publicly available YouTube videos and did not involve human participants, patient data, animal subjects, or identifiable personal information. Therefore, ethics committee approval and informed consent were not required.

Informed Consent: Not applicable.

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

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

Use of AI for Writing Assistance: During the preparation of this manuscript, ChatGPT (OpenAI, GPT-5.5 Pro) was used for language editing, Turkish–English translation support, abstract shortening according to journal word limits, title refinement, and drafting of non-scientific submission documents such as the cover letter.

Author Contributions: Concept – A.Ö., M.O.A.; Design – A.Ö.; Supervision – İ.B.M., G.G.; Data collection and/or processing – İ.B.M., G.G.; Data analysis and/or interpretation – A.Ö.; Literature search – M.O.A.; Writing – A.Ö., G.G.; Critical review – G.G., A.Ö.

Acknowledgments: The authors would like to thank all contributors involved in the evaluation process of the videos included in this study.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Fassnacht M, Tsagarakis S, Terzolo M, Tabarin A, Sahdev A, Newell-Price J, et al. European Society of Endocrinology clinical practice guidelines on the management of adrenal incidentalomas, in collaboration with the European Network for the Study of Adrenal Tumors. Eur J Endocrinol 2023;189:G1–G42. [CrossRef]
2. Brix D, Allolio B, Fenske W, Agha A, Dralle H, Jurowich C, et al. Laparoscopic versus open adrenalectomy for adrenocortical carcinoma: surgical and oncologic outcome in 152 patients. Eur Urol 2010;58:609–15. [CrossRef]
3. Autorino R, Bove P, De Sio M, Miano R, Micali S, Cindolo L, et al. Open versus laparoscopic adrenalectomy for adrenocortical carcinoma: a meta-analysis of surgical and oncological outcomes. Ann Surg Oncol 2016;23:1195–202. [CrossRef]
4. Giordano A, Feroci F, Podda M, Botteri E, Ortenzi M, Montori G, et al. Minimally invasive versus open adrenalectomy for adrenocortical carcinoma: the keys surgical factors influencing the outcomes-a collective overview. Langenbecks Arch Surg 2023;408:256. [CrossRef]
5. Zheng GY, Li HZ, Deng JH, Zhang XB, Wu XC. Open adrenalectomy versus laparoscopic adrenalectomy for adrenocortical carcinoma: a retrospective comparative study on short-term oncologic prognosis. Onco Targets Ther 2018;11:1625–32. [CrossRef]
6. Rapp AK, Healy MG, Charlton ME, Keith JN, Rosenbaum ME, Kapadia MR. YouTube is the most frequently used educational video source for surgical preparation. J Surg Educ 2016;73:1072–6. [CrossRef]

7. Mota P, Carvalho N, Carvalho-Dias E, João Costa M, Correia-Pinto J, Lima E. Video-based surgical learning: improving trainee education and preparation for surgery. J Surg Educ 2018;75:828–35. [CrossRef]
8. Farag M, Bolton D, Lawrentschuk N. Use of YouTube as a resource for surgical education-clarity or confusion. Eur Urol Focus 2020;6:445–9. [CrossRef]
9. Gorgy A, El Hawary H, Galli R, MacDonald M, Barone N, Thibaudeau S. Evaluating the educational quality of surgical YouTube® videos: a systematic review. Health Sci Rev (Oxf) 2023;5:100067. [CrossRef]
10. de'Angelis N, Gavriilidis P, Martínez-Pérez A, Genova P, Notarnicola M, Reitano E, et al. Educational value of surgical videos on YouTube: quality assessment of laparoscopic appendectomy videos by senior surgeons vs. novice trainees. World J Emerg Surg 2019;14:22. [CrossRef]
11. Rouhi AD, Roberson JL, Kindall E, Ghanem YK, Ndong A, Yi WS, et al. What are trainees watching? Assessing the educational quality of online laparoscopic cholecystectomy training videos using the LAP-VEGaS guidelines. Surgery 2023;174:524–8. [CrossRef]
12. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health 1999;53:105–11. [CrossRef]
13. Silberg WM, Lundberg GD, Musacchio RA. Assessing, controlling, and assuring the quality of medical information on the Internet: Caveant lector et viewor--Let the reader and viewer beware. JAMA 1997;277:1244–5. [CrossRef]
14. Celentano V, Smart N, Cahill RA, Spinelli A, Giglio MC, McGrath J, et al. Development and validation of a recommended checklist for assessment of surgical videos quality: the LAParoscopic surgery Video Educational GuidelineS (LAP-VEGaS) video assessment tool. Surg Endosc 2021;35:1362–9. [CrossRef]
15. Rouhi AD, Roberson JL, Kindall E, Ghanem YK, Yi WS, Williams NN, et al. Assessment of YouTube as an online educational tool in teaching laparoscopic Roux-en-Y gastric bypass: A LAP-VEGaS study. Surg Pract Sci 2023;14:100199. [CrossRef]
16. Hashemi SA, Shakiba B, Golshan A, Esmaeil Soofian S, Maghsoudi R. Assessment of the quality of YouTube educational videos on laparoscopic right adrenalectomy and laparoscopic partial nephrectomy surgeries. J Laparoendosc Adv Surg Tech A 2024;34:300–4. [CrossRef]
17. Ahmet A, Gamze K, Rustem M, Sezen KA. Is video-based education an effective method in surgical education? a systematic review. J Surg Educ 2018;75:1150–8. [CrossRef]
18. Green JL, Suresh V, Bittar P, Ledbetter L, Mithani SK, Allori A. The utilization of video technology in surgical education: a systematic review. J Surg Res 2019;235:171–80. [CrossRef]
19. McQueen S, McKinnon V, VanderBeek L, McCarthy C, Sonnadara R. Video-based assessment in surgical education: a scoping review. J Surg Educ 2019;76:1645–54. [CrossRef]
20. Augestad KM, Butt K, Ignjatovic D, Keller DS, Kiran R. Video-based coaching in surgical education: a systematic review and meta-analysis. Surg Endosc 2020;34:521–35. [CrossRef]
21. Srinivasa K, Moir F, Goodyear-Smith F. The role of online videos in teaching procedural skills in postgraduate medical education: a scoping review. J Surg Educ 2022;79:1295–307. [CrossRef]
22. Youssef SC, Aydin A, Canning A, Khan N, Ahmed K, Dasgupta P. Learning surgical skills through video-based education: a systematic review. Surg Innov 2023;30:220–38. [CrossRef]
23. Lima DL, Viscarret V, Velasco J, Lima RNCL, Malcher F. Social media as a tool for surgical education: a qualitative systematic review. Surg Endosc 2022;36:4674–84. [CrossRef]
24. Pakkasjärvi N, Anttila H, Pyhältö K. What are the learning objectives in surgical training - a systematic literature review of the surgical competence framework. BMC Med Educ 2024;24:119. [CrossRef]
25. Yamamoto T, Goto K, Kitano S, Maeshima Y, Yamada T, Azuma Y, et al. Current insights on social media as a tool for the dissemination of research and education in surgery: a narrative review. Surg Today 2024;54:1113–23. [CrossRef]
26. Dantas I, Jorge I, Nicolau A, Vales M, Coutinho C, Rodrigues S, et al. Video-based learning for basic surgical skills - A randomized trial. Surg Open Sci 2025;26:94–106. [CrossRef]

DOI: 10.14744/cm.2026.51422
Compreh Med 2026;18(Suppl. 1):67-73

RESEARCH ARTICLE

Long-Term Anal Fistula Formation After Incision and Drainage of Cryptoglandular Perianal Abscess: A Single-Center Retrospective Cohort Study with 10-Year Follow-up

Emre Özoran[1], Tutku Tüfekçi[2]

[1]Department of General Surgery, Nişantaşı University Faculty of Medicine, İstanbul, Türkiye
[2]Department of General Surgery, Ağrı İbrahim Çeçen University Faculty of Medicine, Ağrı, Türkiye

ABSTRACT

Objective: Perianal abscess (PA) and anal fistula (AF) are considered sequential manifestations of cryptoglandular infection. Long-term data on fistula formation and recurrence after incision and drainage (I&D) remain limited. Most published series use short follow-up periods. This study aimed to determine the long-term AF formation rate (primary outcome) and abscess recurrence rate (secondary outcome) after I&D for PA, using a cohort with a minimum follow-up of 10 years.

Materials and Methods: This single-center retrospective cohort study included all consecutive adult patients who underwent I&D for PA between April 2014 and April 2016. Later events were ascertained by cross-referencing the institutional surgical database (2014–2026) with the Turkish national e-Health record system (e-Nabız), so fistula or recurrence episodes treated at other institutions could be captured. Patients with inflammatory bowel disease, colorectal malignancy, prior perianal surgery, immunosuppression, or age under 18 years were excluded. The study was reported using the STROBE guidelines.

Results: A total of 226 patients were included (145 male [64.2%]; mean age: 37.4 ± 12.0 years). AF developed in 54 patients (23.9%) at a median of 11.0 months (range: 1–105 months) after I&D. Transsphincteric fistulas were the most common (35/54, 64.8%). Abscess recurrence occurred in 53 patients (23.5%) (median: 9.1 months [range: 1.0–99.6 months]). Two outcomes overlapped in 29 patients. The observed fistula rate was consistent with published estimates (25–40%). The wide range of time-to-diagnosis (1–105 months) indicates that short follow-up may underestimate the cumulative incidence.

Conclusion: Approximately one-quarter of patients developed AF, and a further quarter developed recurrent abscesses. Events occurred beyond the early postoperative period. Within the limits of a database-based design, these findings support counseling about the long-term risk of fistula formation and recurrence after I&D for PA.

Keywords: Anal fistula, cryptoglandular, drainage, follow-up studies, perianal abscess, recurrence

How to cite this article: Özoran E, Tüfekçi T. Long-Term Anal Fistula Formation After Incision and Drainage of Cryptoglandular Perianal Abscess: A Single-Center Retrospective Cohort Study with 10-Year Follow-up. Compreh Med 2026;18(Suppl. 1):67-73.

INTRODUCTION

Perianal abscess and anal fistula correspond to the acute and chronic stages, respectively, of cryptoglandular infection that originates from the anal glands located at the dentate line. The cryptoglandular theory continues to serve as the principal pathophysiological framework.[1,2] Obstruction and subsequent infection of these glands lead to the accumulation of pus and the formation of abscesses. Incomplete drainage may subsequently lead to the development of a persistent epithelialized fistulous tract.[2,3]

Incision and drainage (I&D) is widely recognized as the primary surgical intervention for perianal abscess.[4] However, a significant proportion of patients subsequently develop an anal fistula, and another subset experiences recurrent abscess for-

Address for Correspondence: Emre Özoran, Department of General Surgery, Nişantaşı University School of Medicine, İstanbul, Türkiye
E-mail: ozoran@gmail.com, emre.ozoran@nisantasi.edu.tr **ORCID ID:** 0000-0002-9371-6811

Received date: 05.08.2026
Revised date: 25.08.2026
Accepted date: 01.09.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

mation. Reported fistula rates vary widely, from approximately 25% to 40% in most adult series, reflecting differences in study population, follow-up duration, and diagnostic methods.[5,6]

Long-term data are particularly valuable, as fistulas may develop months to years after the index drainage. However, most published series rely on short follow-up periods, and few have systematically cross-referenced hospital records to capture patients who present later, often under a different registration number. This may lead to an underestimation of the true cumulative event rate.

We conducted a single-center retrospective cohort study of patients who underwent I&D for cryptoglandular perianal abscess between April 2014 and April 2016. Later events were ascertained through a cross-referential search of the institutional surgical database using patient name and national identification number, providing a minimum potential observation period of 10 years. The primary outcome was anal fistula formation; the secondary outcome was abscess recurrence.

MATERIALS and METHODS

Study Design and Setting

This single-center, retrospective cohort study included patients who underwent incision and drainage of anal abscesses at the Department of General Surgery, Ağrı State Hospital (subsequently restructured as Ağrı Training and Research Hospital), which served as the referral center for a province with a population of approximately 550,000 during the study period. Ethics approval was granted by the Ethics Committee of Ağrı İbrahim Çeçen University on April 1, 2026 (approval number: 200). The study was conducted in accordance with the principles outlined in the Declaration of Helsinki.[7] As this was a retrospective study based entirely on data obtained from the hospital database, the ethics committee waived the requirement for informed consent. All data were anonymized in accordance with the Turkish Personal Data Protection Law.

Patient Selection

All consecutive patients who underwent I&D for a primary cryptoglandular perianal abscess between April 1, 2014, and April 1, 2016, with the surgery code of "perianal abscess drainage," were included. Inclusion criteria were age equal to or greater than 18 years and I&D as the sole primary procedure. Exclusion criteria were the presence of inflammatory bowel disease, the presence of malignant colorectal disease, immunosuppressive therapy, prior perianal surgery, age under 18 years, and unavailability of follow-up data.

Data Collection and Follow-Up Strategy

Baseline data were extracted from operative records, anesthesia charts, and discharge summaries. Ascertainment of later events did not rely on active clinical follow-up or scheduled outpatient visits. Instead, later fistula and recurrence events were identified retrospectively through a systematic cross-referential search of the complete institutional surgical database (April 2014 to April 2026). Because patients in the Turkish hospital registration system receive a new protocol number at each admission, a matching strategy based on patient name and national identification number was used alongside protocol-number matching, so that repeat surgical encounters for the same patient could be linked regardless of the admission number. To capture events managed outside our institution, each patient's record was additionally verified against the Turkish national e-Health system (e-Nabız), a centralized platform that aggregates diagnosis, procedure, and admission data from all public and private healthcare providers nationwide. This allowed confirmation of each patient's treatment status and identification of fistula or recurrence procedures performed at other centers, substantially reducing the risk of missed events and of misclassifying patients as event-free.

Outcome Definitions

- **Primary outcome:** Anal fistula formation, confirmed clinically or by imaging at any point during the observation period. Fistulas were classified according to the Parks classification.[8] The date of fistula diagnosis was defined as the date of the surgical encounter or imaging study at which the fistula was first documented.
- **Secondary outcome:** Recurrent perianal abscess, defined as a repeat surgical drainage procedure for perianal abscess after the index operation.

The two outcomes were not mutually exclusive and were analyzed both independently and with respect to their overlap. For time-to-event analysis, follow-up began on the date of index drainage. For the primary outcome, patients without a documented fistula were censored at the end of the study observation period (April 1, 2026), yielding a minimum potential follow-up of 10 years and a maximum of 12 years. Time-to-event was recorded in months for both outcomes.

Statistical Analysis

Continuous variables with skewed distributions and time-to-event data were reported as median with range (minimum–maximum). Categorical variables were presented as frequency and percentage. The chi-square test and Fisher's exact test were used to assess associations between categor-

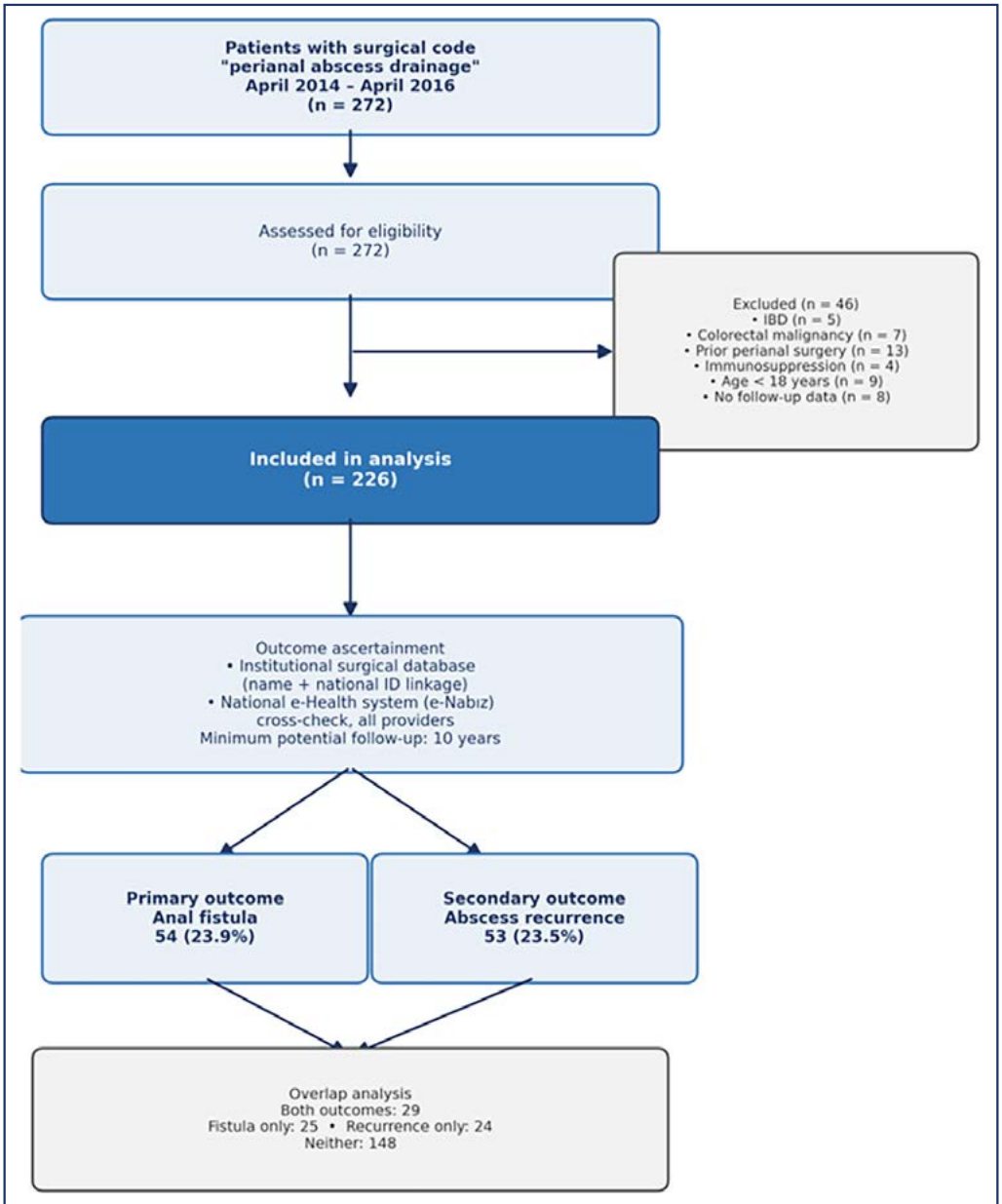


Figure 1. STROBE flow diagram of participant selection and outcome ascertainment. Later events were identified by linking the institutional surgical database with the national e-Health system (e-Nabiz). Exclusion sub-counts are shown

ical variables, with Fisher's exact test applied when expected cell counts were fewer than five.

For time-to-event analysis, the Kaplan-Meier method was used to estimate fistula-free survival, with follow-up measured from the date of index drainage. Patients who developed a fistula were treated as events at the date of diagnosis; all remaining patients were censored at the end of the study observation period (April 1, 2026), since the database search extended to that date and no patient was known to have died or emigrated. Because abscess complexity and other risk factors were not consistently recorded, no multivariable model was fitted, and any subgroup comparison was considered exploratory and hypothesis-generating only.

A two-tailed p-value below 0.05 was considered statistically significant. Analyses were performed using IBM SPSS Statistics version 29.0 (IBM Corp., Armonk, NY, USA). The study was reported in accordance with the STROBE statement for observational studies, and the flow of participants is shown in Figure 1.

Table 1. Baseline patient and operative characteristics (n=226)

Variable	n (%) or mean±SD
Sex	
Male	145 (64.2)
Female	81 (35.8)
Age (years)	
Mean±SD (range)	37.4±12.0 (18-90)
Anesthesia type	
Local	96 (42.5)
Sedation + Local	92 (40.7)
Spinal	31 (13.7)
General	7 (3.1)
Drain placed	39 (17.3)
Penrose	35 (15.5)
Other	4 (1.8)
Intraoperative probing	9 (4.0)
Fistula identified	2 (0.9)
No fistula	7 (3.1)
Perioperative antibiotic prophylaxis	226 (100)
Cefuroxime & Metronidazole	214 (94.7)
Ciprofloxacin+Metronidazole	10 (4.4)
Piperacillin-tazobactam (Tazocin)	2 (0.9)

SD: Standard deviation

Table 2. Primary outcome - anal fistula formation (n=54)

Variable	n	% (of fistula patients)
Fistula formation, overall	54 / 226	23.9% of cohort
Fistula type		
Transsphincteric	35	64.8
Intersphincteric	15	27.8
Extrasphincteric	4	7.4
Confirmation		
Imaging (Pelvic MRI)	49	90.7
Surgical re-exploration	5	9.3
Treatment		
Seton placement	27	50.0
Fistulotomy	20	37.0
Fistulectomy	5	9.3
Flap repair	1	1.9
Conservative management	1	1.9
Time to fistula (months)		
Median (range)		11.0 (1–105)

Table 3. Secondary outcome - abscess recurrence (n=53)

Variable	n	% of cohort
Abscess recurrence, overall	53/226	23.5
Time to recurrence (months) Median (range)		9.1 (1.0–99.6)

RESULTS

Patient Characteristics

During the study period, 272 patients with the surgical code of perianal abscess drainage were assessed for eligibility. Forty-six patients were excluded (inflammatory bowel disease, n=5; colorectal malignancy, n=7; prior perianal surgery, n=13; immunosuppression, n=4; age under 18 years, n=9; no follow-up data, n=8), leaving 226 patients in the final analysis (Fig. 1). Table 1 shows their demographic and operative characteristics.

There were 145 male patients (64.2%) and 81 female patients (35.8%), with a mean age of 37.4 ± 12.0 years (range: 18–90). Anesthesia was local in 96 patients (42.5%), sedation with local in 92 (40.7%), spinal in 31 (13.7%), and general in 7 (3.1%). The procedure consisted of incision over the point of maximal fluctuance, evacuation of pus, and irrigation of the abscess cavity with hydrogen peroxide and/or isotonic saline. A surgical drain was placed at the operating surgeon's discretion when the cavity was large or when dependent drainage was considered necessary; a drain was used in 39 cases (17.3%), most commonly a Penrose drain (35 of 39). Intraoperative probing to identify an internal opening was performed only when an overt fistulous tract was evident at drainage, in 9 patients (4.0%); of these, a fistula was identified in 2. The median follow-up for patients who did not develop a fistula (censored observations) was 122.3 months (10.2 years; range: 99.7–144.9 months), confirming that the great majority of the cohort was observed for at least 10 years.

Primary Outcome: Fistula Formation

Among 226 patients, 54 (23.9%) developed an anal fistula during the observation period (Table 2; Fig. 2). The median time from index drainage to fistula diagnosis was 11.0 months (range: 1–105 months). Transsphincteric fistulas were the most common type, observed in 35 patients (64.8%), followed by intersphincteric fistulas in 15 patients (27.8%) and extrasphincteric fistulas in 4 patients (7.4%).

Fistula diagnosis was confirmed by pelvic MRI in 49 patients (90.7%) and by surgical re-exploration in 5 patients (9.3%). Seton placement was the most frequently performed pro-

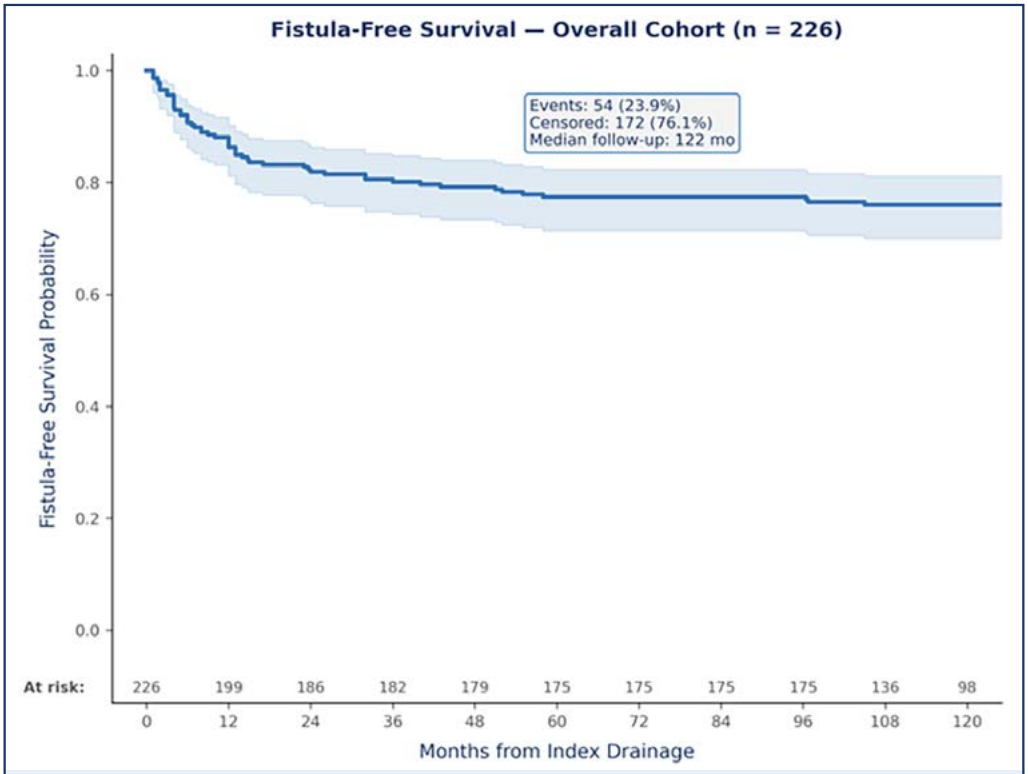


Figure 2. Kaplan-Meier fistula-free survival curve for the overall cohort (n=226). Tick marks indicate censored patients, and the table below the x-axis shows the number at risk at each time point. All patients without a documented fistula were censored at the end of the observation period (April 1, 2026)

cedure (27 patients, 50.0%), followed by fistulotomy (20 patients, 37.0%), fistulectomy (5 patients, 9.3%), advancement flap repair (1 patient, 1.9%), and conservative management (1 patient, 1.9%).

Secondary Outcome: Abscess Recurrence

Perianal abscess recurred in 53 patients (23.5%) during the observation period (Table 3). The median time to first recurrence was 9.1 months (range: 1.0–99.6 months). The primary and secondary outcomes were not mutually exclusive: 29 patients experienced both fistula formation and recurrence, 25 developed a fistula without recurrence, 24 had recurrence without a documented fistula, and 148 experienced neither during the observation period. The cross-referential search using patient name and national identification number was necessary to capture readmission episodes that were registered under new protocol numbers and would otherwise have been missed.

Exploratory Analysis: Anesthesia Type and Fistula Formation

As a hypothesis-generating analysis reported only for completeness, we examined the crude association between anesthesia type and fistula formation. Fistula formation was less frequent in patients drained under local anesthesia than under non-local techniques (12.5% vs. 32.3%). This unadjusted difference must not be interpreted as causal. It is very likely explained by confounding by indication, because smaller and more superficial abscesses are typically drained under local anesthesia, while deeper or more complex abscesses require spinal or general anesthesia; abscess size, depth, and anatomical complexity were not recorded and could not be adjusted for. Recurrence did not differ between groups.

DISCUSSION

In this single-center retrospective cohort, anal fistula developed in 23.9% of patients and recurrent abscess in 23.5% after I&D for cryptoglandular perianal abscess. The administrative observation window spanned April 2014 to April 2026, giving a potential follow-up of 10 to 12 years, and the median follow-up for censored (non-event) patients was 10.2 years. Rather than claiming that these represent the longest available follow-up data, we interpret them as a reminder that a substantial share of adverse events occurs late and would be missed by the short follow-up used in most studies.

The 23.9% fistula rate is consistent with prospective and observational series. In a multicenter observational study, Pearce et al.[5] found that roughly one in five patients developed a fistula within six months of drainage. Hasan et al.[6] reported a fistula incidence of 33.9%, with most cases diagnosed within six months. The American Society of Colon and Rectal Surgeons and the European Society of Coloproctology guidelines similarly describe fistula rates in the range of one-quarter to one-third after simple drainage.[9,10] Our figure sits within this range, reinforcing that a clinically important minority of patients will develop a fistula despite adequate initial drainage.

The median time to fistula formation was 11.0 months, but events were recorded as late as 105 months after drainage. Because most prospective trials use follow-up windows of 6 to 18 months, a meaningful proportion of late fistulas would escape detection in such designs. This is consistent with primary cohort studies that used longer observation: in a multicenter cohort of 475 patients followed for a median of approximately three years, 33.7% developed an anal fistula, while an earlier cohort of 148 patients followed for a mean of 38 months reported a cumulative fistula or recurrent-sepsis rate of 36.5%.[11,12] Our data support counseling patients that fistula risk persists well beyond the first postoperative year.

Transsphincteric fistulas predominated (64.8%). These tracts cross part or all of the external sphincter and are generally more complex to manage than intersphincteric fistulas; the frequent use of seton placement (50.0%) in our cohort reflects a sphincter-preserving approach appropriate to this anatomy.[3] However, because abscess anatomy was not systematically recorded, we cannot determine whether these fistulas arose from deeper index abscesses or were occult at the time of initial drainage.

The 23.5% recurrence rate is consistent with the understanding that simple drainage does not address the internal opening that maintains the cryptoglandular process.[2] The substantial overlap between recurrence and fistula formation (29 patients) supports the view that these outcomes share a common anatomical origin.

All patients received perioperative antibiotic prophylaxis, most commonly cefuroxime with metronidazole; ten patients with recorded allergies received ciprofloxacin with metronidazole, and two with more severe presentations received piperacillin-tazobactam. Because every patient received antibiotics, this cohort has no untreated comparison group and cannot evaluate whether antibiotics influence fistula formation or recurrence. We therefore draw no conclusion regarding antibiotic effectiveness; the antibiotic regimen is reported only to describe institutional practice.

The exploratory finding of a lower fistula rate after local anesthesia is, as noted above, most plausibly explained by confounding by indication and should not be regarded as a clinical signal. We include it only to document the distribution of outcomes across anesthesia groups and to caution against overinterpretation of such associations in retrospective data.

A methodological strength of this study is that ascertainment was not confined to our own institutional records. In addition to cross-referential linkage of surgical records by patient name and national identification number, each patient was verified against the national e-Health system (e-Nabız), which aggregates records from all public and private providers in the country. This national-level cross-check materially reduces the principal weakness of single-center retrospective follow-up, namely the loss of patients who seek later care elsewhere, and increases confidence that the observed event rates approximate the true cumulative incidence in this cohort.

Limitations

This study has important limitations. First, outcomes were ascertained retrospectively rather than through active, scheduled clinical follow-up. Although the national e-Health cross-check substantially reduced the likelihood of missing events treated at other institutions, any event not captured in that system — for example, one managed entirely without a coded procedure — could still be missed; the 10- to 12-year interval therefore represents potential observation time rather than uniformly confirmed active follow-up.

Second, several clinically important risk factors for fistula formation were not consistently recorded, including diabetes mellitus, smoking status, body mass index, abscess size, abscess depth, and anatomical location or complexity. Their absence precluded any multivariable adjustment and means that no risk-factor analysis could be performed; in particular, the lack of abscess classification is a substantial limitation for interpreting fistula risk.

Third, the anesthesia analysis is exploratory and vulnerable to confounding by indication, as discussed above, and should not inform clinical decisions.

Fourth, all patients received antibiotic prophylaxis, so no inference about an antibiotic effect is possible.

Fifth, the single-center design for index treatment and the absence of a comparison group treated with an alternative strategy limit generalizability and preclude causal infer-

ence. These limitations should temper the interpretation of all findings reported here.

CONCLUSION

Within the constraints of a retrospective, single-center, database-based design, this study provides a real-world estimate of the rate of anal fistula formation and abscess recurrence after I&D for cryptoglandular perianal abscess. Anal fistula developed in 23.9% of patients and recurrent abscess in 23.5%, with events occurring as late as 105 months (median: 11.0 months for fistula), indicating that a substantial share of adverse events arises well beyond the early postoperative period. Because ascertainment relied on existing records, these figures should be regarded as a lower-bound estimate.

Two practical implications follow:

1. Prospectively planned, structured databases with active long-term follow-up are needed to define these risks more precisely.
2. Operative and surgical progress notes should be standardized to routinely record abscess size, abscess type, anatomical location, and the surgical technique used, as the absence of these variables was the principal barrier to risk-factor analysis in the present cohort.

Until such data are systematically captured, patients undergoing abscess drainage should be counseled about the meaningful long-term risk of fistula formation and recurrence.

Disclosures

Ethics Committee Approval: The study was approved by the Ağrı İbrahim Çeçen University Ethics Committee (No: 200, Date: 01/04/2026).

Informed Consent: As this was a retrospective study based entirely on data obtained from the hospital database, the ethics committee waived the requirement for informed consent.

Conflict of Interest Statement: The authors declare that they have no conflict of interest.

Funding: The authors declared that this study has received no financial support.

Use of AI for Writing Assistance: AI-assisted tools were used for language editing and formatting; the authors reviewed and take full responsibility for all content.

Author Contributions: Concept – E.Ö.; Design – E.Ö.; Supervision – E.Ö.; Data collection and/or processing – E.Ö., T.T.; Data analysis and/or interpretation – E.Ö.; Literature search – E.Ö., T.T.; Writing – E.Ö.; Critical review – E.Ö., T.T.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Parks AG. Pathogenesis and treatment of fistula-in-ano. Br Med J 1961;1:463–9. [CrossRef]
2. Gosselink MP, van Onkelen RS, Schouten WR. The cryptoglandular theory revisited. Colorectal Dis 2015;17:1041–3. [CrossRef]
3. Włodarczyk M, Włodarczyk J, Sobolewska-Włodarczyk A, Trzciński R, Dziki Ł, Fichna J. Current concepts in the pathogenesis of cryptoglandular perianal fistula. J Int Med Res 2021;49:3000605211001059. [CrossRef]
4. Shi Y, Zhi C, Cheng Y, Zhang C, Zhang Y, Zhan H, et al. A systematic review and meta-analysis of incision and seton drainage in the treatment of high perianal abscess. Ann Palliat Med 2021;10:9830–40. [CrossRef]
5. Pearce L, Newton K, Smith SR, Barrow P, Smith J, Hancock L, et al. Multicentre observational study of outcomes after drainage of acute perianal abscess. Br J Surg 2016;103:1063–8. [CrossRef]
6. Hasan ZAIY, Mohamed B, Alsayegh R, Almarzooq R. Incidence of anal fistula after pyogenic perianal abscess drainage in the Kingdom of Bahrain. Ann Coloproctol 2023;39:169–74. [CrossRef]
7. World Medical Association. World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects. JAMA 2013;310:2191–4. [CrossRef]
8. Parks AG, Gordon PH, Hardcastle JD. A classification of fistula-in-ano. Br J Surg 1976;63:1–12. [CrossRef]
9. Gaertner WB, Burgess PL, Davids JS, Lightner AL, Shogan BD, Sun MY, et al. The American Society of Colon and Rectal Surgeons clinical practice guidelines for the management of anorectal abscess, fistula-in-ano, and rectovaginal fistula. Dis Colon Rectum 2022;65:964–85. [CrossRef]
10. Reza L, Gottgens K, Kleijnen J, Breukink S, Ambe PC, Aigner F, et al. European Society of Coloproctology: guidelines for diagnosis and treatment of cryptoglandular anal fistula. Colorectal Dis 2024;26:145–96. [CrossRef]
11. Skovgaards DM, Perregaard H, Dibbern CB, Nordholm-Carstensen A. Fistula development after anal abscess drainage-a multicentre retrospective cohort study. Int J Colorectal Dis 2023;39:4. [CrossRef]
12. Hamadani A, Haigh PI, Liu IL, Abbas MA. Who is at risk for developing chronic anal fistula or recurrent anal sepsis after initial perianal abscess? Dis Colon Rectum 2009;52:217–21. [CrossRef]

DOI: 10.14744/cm.2026.34976
Compreh Med 2026;18(Suppl. 1):74-83

RESEARCH ARTICLE

Posttraumatic Growth and Depreciation After Childbirth: The Role of Mode of Delivery and Psychological Resilience in a Turkish Postpartum Sample

Berivan Güzelbağ[1,2], Melike Punduk Yılmaz[2], Filiz Yarşılıkal Güleroğlu[1], Ali Çetin[1]

[1]Department of Obstetrics and Gynecology, University of Health Sciences, Haseki Training and Research Hospital, İstanbul, Türkiye
[2]Department of Molecular and Medical Genetics, Biruni University, Institute of Graduate Studies, İstanbul, Türkiye

ABSTRACT

Objective: This study aimed to examine the effects of mode of delivery on posttraumatic growth (PTG), posttraumatic depreciation (PTD), and psychological resilience in postpartum women and to explore their associations with postpartum depression (PPD) and each other.

Materials and Methods: This cross-sectional study included 174 postpartum women attending a tertiary training and research hospital in Istanbul, Türkiye, divided into three groups: vaginal delivery (n=61), elective cesarean section (CS; n=58), and emergency CS (n=55). PPD was assessed using the Edinburgh Postnatal Depression Scale (EPDS), resilience with the Brief Resilience Scale (BRS), and posttraumatic changes with the Posttraumatic Growth and Depreciation Inventory – Expanded Version (PTGDI-X). Data were analyzed using one-way ANOVA with Tukey post-hoc tests, ANCOVA, Pearson correlation, and multiple linear regression.

Results: No significant difference was found in PPD risk across delivery groups. Psychological resilience was significantly lower in the emergency CS group compared with the vaginal and elective CS groups ($p<0.001$). PTG was significantly higher in the emergency CS group than in the elective CS group ($p=0.006$), while PTD was significantly higher in the emergency CS group than in both other groups ($p<0.05$). Emergency CS and BRS independently predicted PTG ($\beta=0.279$ and $\beta=0.295$, respectively), while EPDS, emergency CS, and BRS predicted PTD ($\beta=0.476$, $\beta=0.177$, and $\beta=-0.198$, respectively).

Conclusion: Mode of delivery, particularly emergency CS, is associated with distinct posttraumatic psychological profiles in postpartum women. Routine screening beyond PPD—including PTG, PTD, and resilience—is warranted following operative delivery.

Keywords: Cesarean section, mode of delivery, postpartum depression, posttraumatic depreciation, posttraumatic growth, psychological resilience

How to cite this article: Güzelbağ B, Punduk Yılmaz M, Yarşılıkal Güleroğlu F, Çetin A. Posttraumatic Growth and Depreciation After Childbirth: The Role of Mode of Delivery and Psychological Resilience in a Turkish Postpartum Sample. Compreh Med 2026;18(Suppl. 1):74-83.

INTRODUCTION

PPD is a widely recognized psychiatric condition emerging in the postpartum period, reported to affect around 17% of women globally.[1] PPD is related to serious adverse outcomes for moms and babies, including impaired mother-infant bonding, reduced breastfeeding rates, and long-term neurodevelopmental delays in the child.[1] In Türkiye, a recent meta-analysis of 34 studies involving 10,236 women reported a pooled PPD prevalence of 17.8%, with rates rising to 20.2% during the COVID-19 pandemic.[2] Beyond individual psychosocial risk factors, the mode of delivery has increasingly been recognized as a significant determinant of postpartum psychological outcomes. Türkiye has an extremely high C-section rate, with 60.5% of all deliveries performed by CS in 2022, far exceeding the WHO's suggested threshold of 10–15%.[3–5] Meta-analytic evidence indicates that women who deliver by CS, particularly by emergency CS, are at significantly higher risk for PPD compared with those with

Address for Correspondence: Berivan Güzelbağ, Department of Obstetrics and Gynecology, University of Health Sciences, Sultangazi Haseki Training and Research Hospital, İstanbul, Türkiye; Department of Molecular and Medical Genetics, Biruni University, Institute of Graduate Studies, İstanbul, Türkiye
E-mail: drguzelbag@gmail.com **ORCID ID:** 0000-0002-1141-6381

Received date: 13.05.2026
Revised date: 12.06.2026
Accepted date: 13.06.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

vaginal delivery, with odds ratios ranging from 1.12 to 1.53.[6–8] Importantly, emergency and elective CS may carry distinct psychological implications, as the unplanned and often urgent nature of emergency CS may heighten perceptions of threat and loss of control during childbirth.[9]

Childbirth is a profound life event that, when experienced as traumatic, may initiate significant psychological change. Research suggests that between 20% and 45% of women describe their childbirth experience as traumatic, and posttraumatic stress disorder (PTSD) affects approximately 4% of postpartum women in community samples, rising substantially among those who undergo emergency operative delivery.[10–13] However, the psychological sequelae of traumatic childbirth are not exclusively negative. According to the theoretical model proposed by Tedeschi and Calhoun, individuals who struggle with highly challenging life events may undergo posttraumatic growth (PTG), defined as positive psychological transformation across five domains: personal strength, new possibilities, relating to others, appreciation of life, and spiritual change. Importantly, trauma may simultaneously give rise to posttraumatic depreciation (PTD), reflecting negative changes across these same domains.[14,15] Unlike PTG alone, the concurrent assessment of both PTG and PTD provides a more complete and balanced picture of post-trauma psychological change. The Posttraumatic Growth and Depreciation Inventory – Expanded Version (PTGDI-X), validated across ten countries, including Türkiye, by Senol-Durak and colleagues, is currently the only instrument capable of measuring both constructs simultaneously.[14] In the perinatal context, studies have demonstrated that women who experience traumatic childbirth, particularly those delivered by emergency CS, may report substantial PTG alongside posttraumatic stress symptoms.[16–20]

Psychological resilience, which is generally characterized as the ability to adapt and persist in the face of challenges, has become an important part of postpartum mental health.[21] In the postpartum context, resilience has been shown to buffer against the development of PPD and to facilitate positive psychological adaptation following challenging birth experiences.[22] Evidence from meta-analytic studies indicates a significant positive association between resilience and PTG, suggesting that women with greater psychological resources are better equipped to derive growth from difficult experiences.[22] Conversely, lower resilience has been consistently linked to higher levels of posttraumatic stress and depressive symptoms in the postpartum period.[21] The mode of delivery may further modulate resilience, as the unpredictable and often distressing nature of emergency CS may place greater demands on a woman's adaptive capacities compared with planned vaginal or elective CS delivery. Despite this, the role of psychological resilience as a mediating or moderating variable between mode of delivery and posttraumatic outcomes remains insufficiently explored in the postpartum literature.

To date, no research has examined the simultaneous relationship between postpartum depression (PPD), posttraumatic growth (PTG), posttraumatic depreciation (PTD), and psychological resilience in relation to mode of delivery in a postpartum population. Furthermore, the PTGDI-X, which is the only instrument that can assess posttraumatic growth and depreciation at the same time, has not been used with a Turkish postpartum sample before. Given Türkiye's exceptionally high CS rate and the distinct psychological profiles potentially associated with vaginal, elective CS, and emergency CS deliveries, investigating these outcomes within this context is both timely and clinically relevant. The present study, therefore, aimed to compare PPD, PTG, PTD, and psychological resilience across three delivery groups—vaginal delivery, elective CS, and emergency CS—and to examine the associations and predictors of these outcomes in postpartum women attending a tertiary training and research hospital in Istanbul.

MATERIALS and METHODS

A cross-sectional analytical design was adopted for this study, which was conducted at a tertiary care training and research hospital in Türkiye. Ethical approval was granted by the University of Health Sciences Hamidiye Scientific Research Ethics Committee (Meeting No: 2023/15, Decision No: 15/26, Registration No: 23/425, Date: 11.08.2023), and the study was carried out in accordance with the Declaration of Helsinki. All participants were asked to sign a document confirming their consent before they could join the study.

Postpartum women attending the obstetrics outpatient clinic of a tertiary care hospital in Türkiye were recruited between July 2024 and December 2025. Eligible participants were informed about the study by the research team in the clinic setting. Those who agreed to participate provided written informed consent and were directed to complete an online questionnaire via a QR code linked to a Google Forms platform. All data were collected anonymously without any personal identifying information.

Participants were eligible for inclusion if they (1) were within the first postpartum year and (2) were aged 18 years or

older. Women were excluded if they had (1) a mental or physical disability, (2) an active psychiatric diagnosis or current psychiatric treatment prior to delivery, (3) a clinically significant physical health condition, (4) incomplete questionnaire responses, (5) an inability to read or write in Turkish, (6) perinatal loss (stillbirth or neonatal death), or (7) multiple gestation.

The minimum sample size was calculated using the Raosoft sample size calculator (http://www.raosoft.com/samplesize.html) with a 5% margin of error, a 95% confidence interval, and a 50% response distribution, yielding a minimum of 132 participants. Accounting for a 10% potential dropout rate, a minimum of 145 participants was planned.

Postpartum depression was assessed using the Edinburgh Postnatal Depression Scale.[23,24] A cut-off score of ≥13 indicated depression risk. Psychological resilience was measured with the Brief Resilience Scale.[21,25] Posttraumatic growth and depreciation were evaluated using the Posttraumatic Growth and Depreciation Inventory – Expanded Version, which yields separate PTG and PTD subscale scores ranging from 0 to 125.[14,26] Additional variables, including partner/family support, fear of childbirth, obstetric complications, neonatal intensive care unit (NICU) admission, income level, birth satisfaction, and postpartum duration (days since delivery), were also collected via the questionnaire and hospital records.

Statistical analyses were performed using IBM SPSS Statistics version 31 (IBM Corp., Armonk, NY, USA). Normality of continuous variables was assessed using the Shapiro-Wilk test, and homogeneity of variances was evaluated using Levene's test. Since all scale scores demonstrated a normal distribution and homogeneous variances, one-way analysis of variance (ANOVA) with a Tukey post hoc test was used for group comparisons. Categorical variables were compared using the chi-square (χ^2) test. Pearson correlation analysis was performed to examine associations among scale scores. Multiple linear regression analyses were conducted to identify independent predictors of PTG and PTD. Effect sizes for pairwise group comparisons were calculated using Cohen's d. Analysis of covariance (ANCOVA) was performed to compare scale scores across delivery groups while controlling for age, parity, and BMI. For the regression models, multicollinearity was assessed using the variance inflation factor (VIF), and regression assumptions were examined, including normality of residuals (Shapiro-Wilk test), homoscedasticity (Breusch-Pagan test), and independence of residuals (Durbin-Watson statistic). Statistical significance was set at $p<0.05$.

RESULTS

A total of 174 postpartum women were enrolled in this study within the first postpartum year. The group was organized into three subgroups based on the type of delivery: vaginal delivery (n=61, 35.1%), elective cesarean section (n=58, 33.3%), and emergency cesarean section (n=55, 31.6%). The mean age of all participants was 27.5±4.8 years (range, 18–42). Mean gravidity was 2.2, and mean parity was 1.8. Regarding educational level, 29.3% had primary school education, 27.6% had middle school education, 32.2% had high school education, and 10.9% had university education. The majority of participants were unemployed (61.5%). Most pregnancies were spontaneous (89.1%), and 11.5% of deliveries were preterm. Regarding the additional variables, fear of childbirth and NICU admission differed significantly across delivery groups, both being highest in the emergency CS group ($p=0.029$ and $p=0.002$, respectively). Partner/family support, obstetric complications, income level, birth satisfaction, and postpartum duration did not differ significantly between groups. Detailed sociodemographic and obstetric characteristics stratified by mode of delivery are presented in Table 1.

Cronbach's alpha coefficients were as follows: EPDS $\alpha=0.81$, BRS $\alpha=0.89$, PTG subscale of the PTGDI-X $\alpha=0.89$, and PTD subscale of the PTGDI-X $\alpha=0.89$. All coefficients exceeded the recommended threshold of 0.70.

Mean scale scores stratified by mode of delivery are presented in Table 2. No statistically significant difference was found in EPDS scores across the three groups ($F=0.225$, $P=0.799$). The proportion of women at risk for postpartum depression (EPDS ≥13) was 21.3% in the vaginal delivery group, 29.3% in the elective CS group, and 30.9% in the emergency CS group.

Statistically significant group differences were observed for BRS ($F=14.835$, $p<0.001$), PTG ($F=4.895$, $p=0.009$), and PTD ($F=7.033$, $p=0.001$). Post hoc Tukey comparisons revealed that BRS scores differed significantly between all three groups (vaginal > elective > emergency; all $p<0.05$), with the largest effect observed between vaginal delivery and emergency CS (Cohen's $d=1.04$). For PTG, the emergency CS group scored significantly higher than the elective CS group ($p=0.006$, Cohen's $d=0.62$), while vaginal delivery did not differ significantly from either CS group. For PTD, the emergency CS group scored significantly higher than both the vaginal delivery ($p=0.001$, Cohen's $d=0.63$) and elective CS groups ($p=0.013$, Cohen's $d=0.54$), with no significant difference between the vaginal and elective CS groups.

Table 1. Sociodemographic and obstetric characteristics by mode of delivery

Variable	Vaginal (n=61)	Elective CS (n=58)	Emergency CS (n=55)	p
Continuous variables, mean±SD				
Age (years)	25.7±4.3	29.1±4.8	27.9±4.8	**<0.001**
Gravidity	2.0±0.8	2.6±0.9	2.1±0.8	**<0.001**
Parity	1.6±0.6	2.0±0.9	1.9±0.7	**0.013**
Postpartum duration (days)	110.0±99.2	79.7±85.9	92.9±89.0	0.198
Categorical variables, n (%)				
Educational level				0.183
Primary school	23 (37.7)	14 (24.1)	14 (25.5)	
Middle school	16 (26.2)	16 (27.6)	16 (29.1)	
High school	14 (23.0)	25 (43.1)	17 (30.9)	
University	8 (13.1)	3 (5.2)	8 (14.5)	
Employment status				0.638
Employed	23 (37.7)	25 (43.1)	19 (34.5)	
Unemployed	38 (62.3)	33 (56.9)	36 (65.5)	
Income level				0.126
Low	35 (57.4)	27 (46.6)	33 (60.0)	
Middle	26 (42.6)	26 (44.8)	18 (32.7)	
High	0 (0.0)	5 (8.6)	4 (7.3)	
BMI group				**<0.001**
Normal (<25)	34 (55.7)	19 (32.8)	19 (34.5)	
Overweight (25–29.9)	11 (18.0)	18 (31.0)	31 (56.4)	
Obese (≥30)	16 (26.2)	21 (36.2)	5 (9.1)	
Conception type				0.541
Spontaneous	56 (91.8)	52 (89.7)	47 (85.5)	
Assisted	5 (8.2)	6 (10.3)	8 (14.5)	
Gestational age				**0.037**
Term (≥37 weeks)	58 (95.1)	52 (89.7)	44 (80.0)	
Preterm (<37 weeks)	3 (4.9)	6 (10.3)	11 (20.0)	
Gestational weight gain				**0.022**
<8 kg	3 (4.9)	8 (13.8)	5 (9.1)	
8–12 kg	17 (27.9)	15 (25.9)	29 (52.7)	
13–18 kg	26 (42.6)	23 (39.7)	15 (27.3)	
>18 kg	15 (24.6)	12 (20.7)	6 (10.9)	
Partner/family support				0.323
Insufficient	10 (16.4)	18 (31.0)	13 (23.6)	
Moderate	22 (36.1)	21 (36.2)	22 (40.0)	
Sufficient	29 (47.5)	19 (32.8)	20 (36.4)	
Fear of childbirth				**0.029**
Low	18 (29.5)	16 (27.6)	12 (21.8)	
Moderate	31 (50.8)	20 (34.5)	17 (30.9)	
High	12 (19.7)	22 (37.9)	26 (47.3)	

Table 1. Cont.

Variable	Vaginal (n=61)	Elective CS (n=58)	Emergency CS (n=55)	p
Obstetric complication				0.063
No	50 (82.0)	45 (77.6)	35 (63.6)	
Yes	11 (18.0)	13 (22.4)	20 (36.4)	
NICU admission				**0.002**
No	57 (93.4)	51 (87.9)	39 (70.9)	
Yes	4 (6.6)	7 (12.1)	16 (29.1)	
Birth satisfaction				0.353
Low	20 (32.8)	12 (20.7)	20 (36.4)	
Moderate	25 (41.0)	24 (41.4)	20 (36.4)	
High	16 (26.2)	22 (37.9)	15 (27.3)	

One-Way ANOVA was used for continuous variables and Chi-square (χ^2) test for categorical variables. Bold P values indicate statistical significance (p<0.05). CS: Cesarean section; SD: Standard deviation; BMI: Body mass index; NICU: Neonatal intensive care unit

Table 2. Scale scores by mode of delivery (mean±SD)

Scale	Total (n=174)	Vaginal (n=61)	Elective CS (n=58)	Emergency CS (n=55)	F	p
EPDS (0–30)	9.68±4.05	9.41±3.92	9.78±4.20	9.89±4.07	0.225	0.799
≥13, n (%)	47 (27.0)	13 (21.3)	17 (29.3)	17 (30.9)	–	–
BRS (6–30)	19.64±4.98	21.90±4.21[a]	19.57±4.85[b]	17.22±4.82[c]	14.835	**<0.001**
PTG (0–125)	64.31±21.54	65.03±23.20[ab]	57.91±19.03[a]	70.25±20.65[b]	4.895	**0.009**
PTD (0–125)	38.10±20.35	33.18±19.44[a]	35.60±17.28[a]	46.20±22.15[b]	7.033	**0.001**

One-Way ANOVA with Tukey post-hoc comparisons. Different superscript letters (a, b, c) indicate statistically significant pairwise differences (p<0.05). Bold p values indicate statistical significance. EPDS: Edinburgh postnatal depression scale; BRS: Brief resilience scale; PTG: Posttraumatic growth; PTD: Posttraumatic depreciation; CS: Cesarean section; SD: Standard deviation

Group differences in BRS, PTG, and PTD scores are illustrated in Figure 1.

Pearson correlation analysis revealed a strong negative correlation between EPDS and BRS (r=−0.600, p<0.001) and a strong positive correlation between EPDS and PTD (r=0.593, p<0.001). BRS was negatively correlated with PTD (r=-0.541, p<0.001) and weakly positively correlated with PTG (r=0.152, p=0.045). No statistically significant correlation was found between EPDS and PTG (r=-0.100, p=0.190) or between PTG and PTD (r=-0.065, p=0.392). The full correlation matrix is presented in Table 3.

The identification of independent predictors of PTG and PTD was the objective of the multiple linear regression models. Results are presented in Table 4. In the PTG model (R^2=0.136, adjusted R^2=0.083, F=2.562, p=0.007), emergency CS (β=0.279, p=0.005) and BRS (β=0.295, p=0.006) were significant positive predictors of PTG. In the PTD model (R^2=0.444, adjusted R^2=0.409, F=12.996, p<0.001), EPDS was the strongest predictor of PTD (β=0.476, p<0.001), followed by emergency CS (β=0.177, p=0.025) as a positive predictor and BRS (β=-0.198, p=0.021) as a negative predictor. Elective CS, parity, educational level, and conception type did not significantly predict either PTG or PTD. Postpartum duration was included as a covariate in both regression models but did not significantly predict either PTG (β=0.031, p=0.696) or PTD (β=0.045, p=0.479). Multicollinearity was not a concern, as all VIF values were below 2.2. All regression assumptions were satisfied for both models (Shapiro-Wilk p>0.05; Breusch-Pagan

Table 3. Pearson correlation matrix of scale scores

Scale	1. EPDS	2. BRS	3. PTG	4. PTD
1. EPDS	1			
2. BRS	**-0.600****	1		
3. PTG	-0.100	0.152*	1	
4. PTD	**0.593****	**-0.541****	-0.065	1

*: p<0.05; **: p<0.01. EPDS: Edinburgh postnatal depression scale; BRS: Brief resilience scale; PTG: Posttraumatic growth; PTD: Posttraumatic depreciation

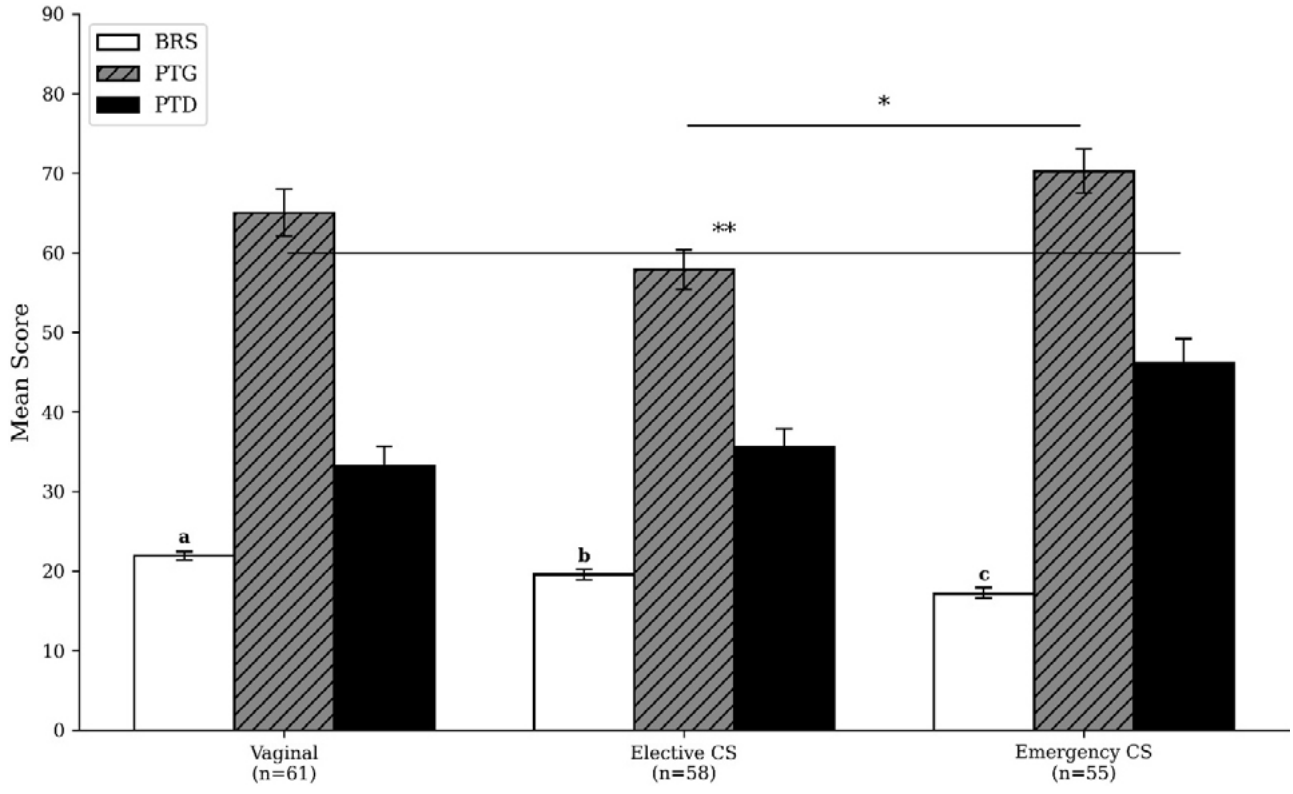


Figure 1. Scale scores by mode of delivery

*Error bars represent SEM. Tukey post-hoc comparisons: *: p<0.05; **: p<0.01. BRS: Brief resilience scale; PTG: Posttraumatic growth; PTD: Posttraumatic depreciation*

p>0.05; Durbin-Watson, 1.83–1.93). In ANCOVA adjusting for age, parity, and BMI, group differences remained significant for BRS (p<0.001), PTG (p=0.011), and PTD (p=0.001), whereas EPDS remained non-significant (p=0.502).

DISCUSSION

In the present study, PPD risk scores, as measured by the EPDS, did not differ significantly across the three delivery groups. This finding is consistent with several previous studies reporting no significant association between mode of delivery and PPD severity when assessed by the EPDS at a cut-off of ≥13.[7,8] Although meta-analytic evidence suggests that CS, particularly emergency CS, is associated with a modestly elevated risk of mild PPD, the association with clinically significant PPD defined by a threshold score is less consistent.[6,7] One plausible explanation is that PPD is a multifactorial condition influenced by a broad range of psychosocial, hormonal, and relational factors and that the mode of delivery may represent only one contributing variable among many.[1] Notably, the proportion of women at risk for PPD (EPDS ≥13) was numerically higher in the emergency CS group (30.9%) compared with the vaginal delivery group (21.3%), a pattern that, while not reaching statistical significance, may carry clinical relevance. Indeed, a difference of approximately 10 percentage points in PPD risk between the vaginal delivery and emergency CS groups (21.3% vs. 30.9%) could potentially reach statistical significance in a larger, adequately powered sample. Future studies with prospective designs and sufficient statistical power to detect group differences in PPD risk are therefore warranted.

Psychological resilience scores differed significantly across delivery groups, with vaginal delivery associated with the highest resilience and emergency CS with the lowest.[22] This finding suggests that the mode of delivery may have a meaningful impact on women's capacity to recover from the challenges of childbirth. The unplanned and often distressing nature of emergency CS, characterized by time pressure, loss of control, and heightened fear for maternal and fetal well-being, may deplete psychological resources and undermine resilience in the postpartum period.[9,12] Interestingly, elective CS was also associated with lower resilience com-

Table 4. Multiple linear regression models for PTG and PTD

Variable	PTG model					PTD model				
	B	SE	β	t	p	B	SE	β	t	p
Mode of delivery (ref: Vaginal)										
Elective CS	-2.036	4.276	-0.045	-0.476	0.635	-0.708	3.241	-0.016	-0.218	0.827
Emergency CS	**12.903**	4.516	**0.279**	2.857	**0.005***	**7.727**	3.423	**0.177**	2.257	**0.025***
Scale Scores										
EPDS	0.226	0.510	0.042	0.443	0.658	**2.394**	0.386	**0.476**	6.196	**<0.001****
BRS	**1.275**	0.457	**0.295**	2.794	**0.006****	**-0.808**	0.346	**-0.198**	-2.336	**0.021***
Obstetric characteristics										
Parity	-0.407	2.292	-0.014	-0.178	0.859	0.323	1.737	0.012	0.186	0.853
Education level	-1.484	1.723	-0.069	-0.861	0.390	-0.167	1.306	-0.008	-0.128	0.899
Conception type	2.807	5.436	0.041	0.516	0.606	5.017	4.120	0.077	1.218	0.225
Age	-0.675	0.358	-0.151	-1.886	0.061	0.058	0.271	0.014	0.214	0.830
BMI	2.145	2.039	0.079	1.052	0.294	0.799	1.546	0.031	0.517	0.606
Postpartum duration	0.007	0.018	0.031	0.392	0.696	0.010	0.014	0.045	0.709	0.479
R^2			**0.136**					**0.444**		
Adjusted R^2			**0.083**					**0.409**		
F (df)		2.562 (10,163)			**0.007**		12.996 (10,163)			**<0.001**

Vaginal delivery was used as the reference category. *: p<0.05; **: p<0.01. B: Unstandardized coefficient; SE: Standard error; β: Standardized coefficient; CS: Cesarean Section; EPDS: Edinburgh postnatal depression scale; BRS: Brief resilience scale; PTG: Posttraumatic growth; PTD: Posttraumatic depreciation; BMI: Body mass index

pared with vaginal delivery, which may reflect the psychological burden of surgical recovery and altered birth expectations even in planned operative deliveries. These results are consistent with emerging evidence suggesting that the subjective experience of childbirth, rather than the objective clinical outcome alone, plays a central role in shaping postpartum psychological resilience.[11] To our knowledge, this is the first study to directly compare psychological resilience across vaginal, elective CS, and emergency CS groups using the BRS in a Turkish postpartum sample, and these findings highlight the need for targeted resilience-building interventions for women undergoing operative delivery.

Posttraumatic growth scores differed significantly across delivery groups, with the emergency CS group reporting the highest PTG and the elective CS group reporting the lowest. Post hoc analysis revealed that the difference was statistically significant only between the elective CS and emergency CS groups (p=0.006), while the vaginal delivery group did not differ significantly from either CS group. These findings are partially consistent with previous research demonstrating that more challenging and unexpected birth experiences may paradoxically serve as a catalyst for psychological growth.[16,19,20] The higher PTG observed in the emergency CS group may reflect the greater degree of cognitive disruption and core belief challenge inherent to unplanned operative delivery, which, according to the theoretical model of Tedeschi and Calhoun, is a prerequisite for the growth process. Notably, women who deliver by CS have been reported to experience higher PTG compared with those with vaginal delivery in other cultural contexts, lending cross-cultural support to the present findings.[15,27] The comparatively lower PTG in the elective CS group may be explained by the planned and anticipated nature of this delivery mode, which may afford less opportunity for the kind of profound psychological struggle that drives posttraumatic growth.[17] Notably, the absence of a significant difference between vaginal delivery and emergency CS groups suggests that growth may occur across a range of birth experiences and that the planned versus unplanned distinction may be more determinative than the route of delivery per se.[18]

Posttraumatic depreciation scores were significantly higher in the emergency CS group compared with both the vaginal delivery and elective CS groups, indicating that unplanned operative delivery is associated not only with greater growth potential but also with more pronounced negative psychological change. This finding underscores the dual nature of post-trauma psychological responses, whereby PTG and

PTD may coexist and operate as independent constructs rather than opposing ends of a single continuum.[14] The elevated PTD in the emergency CS group is consistent with evidence that unplanned CS is associated with heightened posttraumatic stress symptoms, loss of sense of control, and negative appraisal of the birth experience.[9,10,12,28] Importantly, the absence of a significant PTD difference between the vaginal delivery and elective CS groups suggests that it is specifically the unplanned and potentially traumatic nature of emergency CS, rather than the surgical route itself, that drives posttraumatic depreciation. However, it should be acknowledged that emergency CS is frequently performed in the context of acute obstetric complications such as fetal distress, labor arrest, or placental abruption, which may themselves independently contribute to adverse psychological outcomes. Future studies should systematically record emergency CS indications to disentangle the psychological impact of the procedure itself from that of the underlying obstetric emergency. These results highlight the clinical importance of distinguishing between elective and emergency CS when assessing postpartum psychological outcomes and support the utility of the PTGDI-X as a sensitive instrument for capturing the full spectrum of post-trauma change in postpartum women.[14] Consistent with this interpretation, both fear of childbirth and NICU admission were also significantly more frequent in the emergency CS group, suggesting that women undergoing emergency CS experienced a more distressing perinatal course. This accumulation of stressors may partly explain the elevated posttraumatic depreciation observed in this group.

Pearson correlation analysis revealed several significant associations among the study variables. The strong negative correlation between EPDS and BRS scores (r=-0.600) indicates that women with higher levels of depressive symptoms tend to report lower psychological resilience, a finding consistent with the broader literature linking resilience as a protective factor against depression and adverse psychological outcomes in clinical populations.[22] The strong positive correlation between EPDS and PTD scores (r=0.593) suggests that postpartum depressive symptoms and posttraumatic depreciation share considerable common variance, potentially reflecting overlapping cognitive and emotional processes such as negative self-appraisal, hopelessness, and perceived loss of personal competence following childbirth. Similarly, the significant negative correlation between BRS and PTD (r=-0.541) indicates that women with lower resilience are more vulnerable to posttraumatic depreciation, reinforcing the role of psychological resources in buffering against negative post-trauma change.[22] In contrast, PTG and PTD were not significantly correlated (r=-0.065), supporting the theoretical position that these two constructs are independent dimensions of post-trauma change rather than opposite poles of a single continuum.[14,29] This finding has important methodological implications, as it validates the decision to assess PTG and PTD simultaneously using the PTGDI-X rather than relying on PTG measures alone. Additionally, BRS was weakly but significantly positively correlated with PTG (r=0.152, p=0.045), suggesting that higher resilience is modestly associated with greater posttraumatic growth.

Multiple linear regression analyses identified emergency CS and psychological resilience as significant independent predictors of PTG, together explaining 13.5% of its variance. Although age approached but did not reach statistical significance (β=-0.156, p=0.051), the negative trend suggests that younger women may be more prone to posttraumatic growth following childbirth, consistent with previous literature. Women who delivered by emergency CS reported higher PTG (β=0.279, p=0.005), while higher resilience was also associated with greater PTG (β=0.295, p=0.006). These findings suggest that both the nature of the birth experience and individual psychological resources independently contribute to posttraumatic growth in the postpartum period, consistent with theoretical models emphasizing the interplay between event-related challenge and personal strengths in driving PTG.[15,17] For PTD, the regression model explained a substantially larger proportion of variance (R^2=0.444), with three significant predictors identified: EPDS scores (β=0.476, p<0.001), emergency CS (β=0.177, p=0.025), and BRS scores (β=-0.198, p=0.021). The dominant role of depressive symptoms in predicting PTD suggests that posttraumatic depreciation and postpartum depression may share common psychological pathways, including negative cognitive appraisal and emotional dysregulation. The independent contribution of emergency CS to both PTG and PTD underscores the psychological complexity of unplanned operative delivery, simultaneously increasing the risk of negative change while also creating conditions conducive to growth. The protective role of resilience against PTD, alongside its facilitative role for PTG, further highlights psychological resilience as a key intervention target in postpartum care following operative delivery. Importantly, these group differences remained significant after adjusting for age, parity, and BMI in ANCOVA, indicating that the observed effects were not attributable to demographic differences between delivery groups. Moreover, the effect sizes for the observed group differences ranged from moderate to large, supporting their clinical relevance.

Several limitations of the present study should be acknowledged. First, the cross-sectional design precludes any causal inference regarding the relationship between mode of delivery and the psychological outcomes assessed. Longitudinal studies are needed to elucidate the temporal trajectory of PPD, PTG, PTD, and resilience following different delivery modes. Second, data were collected via self-report questionnaires, which are subject to recall bias and social desirability effects, particularly in the context of sensitive postpartum experiences. Third, the study was conducted at a single tertiary training and research hospital in Istanbul, which may limit the generalizability of findings to other geographic regions and healthcare settings in Türkiye. Fourth, although women with active psychiatric diagnoses prior to delivery were excluded, other potential confounders—including posttraumatic stress symptoms and pregnancy intention—were not assessed and therefore could not be controlled for in the regression analyses. This may partly explain the relatively modest variance explained by the PTG model (R^2=0.136), as posttraumatic growth is known to be influenced by a broad range of cognitive, social, and event-specific factors that extend beyond mode of delivery and resilience alone. Fifth, although the time elapsed since delivery varied across participants within the first postpartum year, postpartum duration was included as a covariate in the regression models and did not significantly predict PTG or PTD, suggesting that this heterogeneity did not substantially confound the main findings. Nonetheless, the cross-sectional assessment at varying postpartum time points precludes conclusions about the temporal evolution of these constructs. Finally, emergency CS indications were not systematically recorded in a standardized format; as operative notes were documented by different surgeons, varied considerably across cases, and frequently included multiple concurrent indications, retrospective categorization was not feasible. Future studies should prospectively record emergency CS indications using a standardized classification system to disentangle the psychological effects of the procedure itself from those of the underlying obstetric emergency.

CONCLUSION

In conclusion, the present study demonstrates that mode of delivery is differentially associated with psychological resilience, posttraumatic growth, and posttraumatic depreciation in the postpartum period, with emergency CS emerging as the delivery mode most strongly linked to both negative and positive post-trauma psychological change. Although PPD risk did not differ significantly across delivery groups, emergency CS was independently associated with higher PTD and higher PTG, while lower psychological resilience predicted both greater PTD and lesser PTG. These findings underscore the psychological complexity of operative delivery and highlight the importance of routine postpartum psychological screening that extends beyond depression to include assessment of posttraumatic growth, depreciation, and resilience. Clinically, women who undergo emergency CS may benefit from targeted psychological support interventions in the early postpartum period aimed at strengthening resilience and facilitating adaptive post-trauma processing.[30] Future studies with longitudinal designs, larger and more diverse samples, and objective measures of birth experience are warranted to further elucidate the pathways linking mode of delivery to postpartum psychological outcomes.

Disclosures

Ethics Committee Approval: The study was approved by the University of Health Sciences Hamidiye Scientific Research Ethics Committee (No: 15/26, Date: 11/08/2023).

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

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

Funding: No funding was received for this study.

Use of AI for Writing Assistance: Artificial intelligence tools were used solely for language editing and proofreading assistance. All scientific content, data interpretation, and conclusions are the sole responsibility of the authors.

Author Contributions: Concept – B.G., M.P.Y.; Design – B.G., M.P.Y.; Supervision – F.Y.G., A.Ç.; Funding – B.G., M.P.Y.; Materials – B.G., M.P.Y.; Data collection and/or processing – B.G., M.P.Y.; Data analysis and/or interpretation – B.G.; Literature search – B.G., M.P.Y.; Writing – B.G.; Critical review – F.Y.G., A.Ç.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Wang Z, Liu J, Shuai H, Cai Z, Fu X, Liu Y, et al. Mapping global prevalence of depression among postpartum women. Transl Psychiatry 2021;11(1):543. [CrossRef]
2. Karaçam Z, Ekin P, Şaraldı HB. Comparison of the prevalence of probably postpartum depression before and during the covid-19 pandemic in Turkey: a systematic review and meta-analysis. Soc Psychiatry Psychiatr Epidemiol 2025;60:2469–85. [CrossRef]
3. Betran AP, Ye J, Moller AB, Souza JP, Zhang J. Trends and projections of caesarean section rates: global and regional estimates. BMJ Glob Health 2021;6:e005671. [CrossRef]
4. Birinci Ş, Parpucu ÜM. When a caesarean section is necessary: Analysis of cesarean sections performed in the Republic of Turkey in 2022 in accordance with the World Health Organization Multi-Country Research Guidelines. Turk J Obstet Gynecol 2023;20(3):184–90. [CrossRef]

5. Ulgu MM, Birinci S, Altun Ensari T, Gözükara MG. Cesarean section rates in Turkey 2018-2023: Overview of national data by using Robson ten group classification system. Turk J Obstet Gynecol 2023;20:191–8. [CrossRef]
6. Ning J, Deng J, Li S, Lu C, Zeng P. Meta-analysis of association between caesarean section and postpartum depression risk. Front Psychiatry 2024;15:1361604. [CrossRef]
7. Sun L, Wang S, Li XQ. Association between mode of delivery and postpartum depression: A systematic review and network meta-analysis. Aust N Z J Psychiatry 2021;55:588–601. [CrossRef]
8. Xu H, Ding Y, Ma Y, Xin X, Zhang D. Cesarean section and risk of postpartum depression: A meta-analysis. J Psychosom Res 2017;97:118–26. [CrossRef]
9. Dekel S, Ein-Dor T, Berman Z, Barsoumian IS, Agarwal S, Pitman RK. Delivery mode is associated with maternal mental health following childbirth. Arch Womens Ment Health 2019;22:817–24. [CrossRef]
10. Heyne CS, Kazmierczak M, Souday R, Horesh D, Lambregtse-van den Berg M, Weigl T, et al. Prevalence and risk factors of birth-related posttraumatic stress among parents: A comparative systematic review and meta-analysis. Clin Psychol Rev 2022;94:102157. [CrossRef]
11. Horsch A, Garthus-Niegel S, Ayers S, Chandra P, Hartmann K, Vaisbuch E, et al. Childbirth-related posttraumatic stress disorder: definition, risk factors, pathophysiology, diagnosis, prevention, and treatment. Am J Obstet Gynecol 2024;230:S1116–s27. [CrossRef]
12. Hüner B, Friedl T, Schütze S, Polasik A, Janni W, Reister F. Post-traumatic stress syndromes following childbirth influenced by birth mode-is an emergency cesarean section worst? Arch Gynecol Obstet 2024;309:2439–46. [CrossRef]
13. Yildiz PD, Ayers S, Phillips L. The prevalence of posttraumatic stress disorder in pregnancy and after birth: A systematic review and meta-analysis. J Affect Disord 2017;208:634–45. [CrossRef]
14. Taku K, Tedeschi RG, Shakespeare-Finch J, Krosch D, David G, Kehl D, et al. Posttraumatic growth (PTG) and posttraumatic depreciation (PTD) across ten countries: Global validation of the PTG-PTD theoretical model. Pers Individ Dif 2021;169:110222. [CrossRef]
15. Tedeschi RG, Calhoun LG. TARGET ARTICLE: Posttraumatic growth: Conceptual foundations and empirical evidence. Psychol Inq 2004;15:1–18. [CrossRef]
16. Beck CT, Watson S. Posttraumatic Growth After Birth Trauma: "I Was Broken, Now I Am Unbreakable". MCN Am J Matern Child Nurs 2016;41:264–71. [CrossRef]
17. Brandão T, Brandão S, Prata AP, Silva R, Abreu W, Riklikiene O, et al. Posttraumatic growth in women after a childbirth experience: The influence of individual characteristics and intrusive and deliberate rumination. Psychol Trauma 2025;17:1076–85. [CrossRef]
18. Brandão T, Brites R, Nunes O, Hipólito J. Posttraumatic growth after childbirth in women: A systematic review. J Clin Psychol Med Settings 2020;27(2):318–30. [CrossRef]
19. Ketley R, Darwin Z, Masterson C, McGowan L. Women's experience of post-traumatic growth following a traumatic birth: an interpretive phenomenological analysis. J Reprod Infant Psychol 2024;42:126–37. [CrossRef]
20. Sawyer A, Ayers S. Post-traumatic growth in women after childbirth. Psychol Health 2009;24:457–71. [CrossRef]
21. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008;15:194–200. [CrossRef]
22. Wan X, Huang H, Peng Q, Yu NX, Zhang Y, Ding Y, et al. A meta-analysis on the relationship between posttraumatic growth and resilience in people with breast cancer. Nurs Open 2023;10(5):2734–45. [CrossRef]
23. Cox JL, Holden JM, Sagovsky R. Detection of postnatal depression. Development of the 10-item Edinburgh Postnatal Depression Scale. Br J Psychiatry 1987;150:782–6. [CrossRef]
24. Engindeniz A, Küey L, Kültür S. Edinburgh doğum sonrası depresyon ölçeği Türkçe formu geçerlilik ve güvenilirlik çalışması. Bahar Sempozyumları 1996;1:51–2.
25. Doğan T. Adaptation of the brief resilience scale into Turkish: A validity and reliability study. J Happiness Well-Being 2015;3:93–102. [Turkish]
26. Tedeschi RG, Cann A, Taku K, Senol-Durak E, Calhoun LG. The Posttraumatic Growth Inventory: A Revision Integrating Existential and Spiritual Change. J Trauma Stress 2017;30:11–8. [CrossRef]
27. Suzuki A, Iriyama S, Obayashi Y. Validity and reliability of the Posttraumatic Growth Inventory among Japanese women with cesarean section. Nagoya J Med Sci 2023;85:745–57.
28. Carter J, Bick D, Gallacher D, Chang YS. Mode of birth and development of maternal postnatal post-traumatic stress disorder: A mixed-methods systematic review and meta-analysis. Birth 2022;49:616–27. [CrossRef]
29. Ho SMY, Cheng CT, Shih SM, Taku K, Tedeschi RG. The Chinese version of Posttraumatic Growth and Depreciation Inventory-Expanded version (PTGDI-X) for cancer survivors. Support Care Cancer 2022;30:237–49. [CrossRef]
30. Dekel S, Papadakis JE, Quagliarini B, Pham CT, Pacheco-Barrios K, Hughes F, et al. Preventing posttraumatic stress disorder following childbirth: a systematic review and meta-analysis. Am J Obstet Gynecol 2024;230:610–41.e14. [CrossRef]

DOI: 10.14744/cm.2026.94824
Compreh Med 2026;18(Suppl. 1):84-93

RESEARCH ARTICLE

Integration of RES-Q-Compatible Stroke Documentation into the Hospital Information System: Evaluation of Stroke Center Indicators

Hatice Yelda Yıldız[1], Umut Esen[2], Yavuz Bekmezci[1], Neslihan Oral[3], Nihal Ünlütürk[3], Ferda Kaya Zaman[2], Tülay Özer[1], Şerif Köksal[4], Yakup Krespi[1]

[1]Department of Neurology Stroke Center, İstinye University Faculty of Medicine, İstanbul, Türkiye
[2]Department of Family Medicine, İstinye University Faculty of Medicine, İstanbul, Türkiye
[3]Department of Quality, İstinye University Faculty of Medicine, İstanbul, Türkiye
[4]Medical Directorate, MedicalPark Hospitals Group, İstanbul, Türkiye

ABSTRACT

Objective: To integrate selected RES-Q-compatible quality indicators into the Hospital Information System (HIS) of a comprehensive stroke center and evaluate their impact on the completeness and consistency of structured clinical documentation.

Materials and Methods: A before-and-after observational study compared documentation completeness before and after RES-Q-compatible HIS integration. The pre-integration period covered November 2022–October 2023, and the post-integration period covered January–December 2024. November–December 2023 was considered the implementation phase and excluded from comparative analyses. The study was approved by the Institutional Review Board of Istinye University Human Research Ethics Committee (approval date: November 22, 2024; protocol number: 24-272) and was conducted in accordance with the principles of the Declaration of Helsinki. Structured documentation templates based on RES-Q quality indicators were incorporated into the HIS workflow.

Results: Documentation completeness improved significantly across almost all evaluated domains following HIS integration. Recording of symptom-onset time increased from 57.5% to 89.4%, and wake-up stroke from 46.6% to 85.3% (both p<0.001). Documentation of hypertension, coronary artery disease, diabetes mellitus, and atrial fibrillation increased from 78.8% to 93.6%, 61.1% to 92.6%, 67.0% to 91.9%, and 64.8% to 86.6%, respectively (all p<0.001). NIHSS documentation on admission improved from 47.9% to 91.9%, while INR documentation increased from 23.4% to 42.6% (both p<0.001). Overall, missing data were substantially reduced.

Conclusion: Integration of RES-Q-compatible structured documentation into the HIS significantly improved documentation completeness, data consistency, and workflow standardization in stroke care. Embedding structured data collection into routine clinical practice provides a sustainable framework for quality improvement and future registry-based performance evaluation.

Keywords: Data accuracy, electronic health records, hospital information systems, RES-Q, quality improvement, quality indicators, stroke, stroke registry

How to cite this article: Yıldız HY, Esen U, Bekmezci Y, Oral N, Ünlütürk N, Kaya Zaman F, et al. Integration of RES-Q-Compatible Stroke Documentation into the Hospital Information System: Evaluation of Stroke Center Indicators. Compreh Med 2026;18(Suppl. 1):84-93.

INTRODUCTION

Stroke remains a major global public-health concern and a leading cause of mortality and long-term disability worldwide.[1] Its burden is projected to increase substantially over the coming decades, highlighting the need for systematic approaches to monitor and improve the quality of stroke care.[2] International organizations, including the European Stroke Organisation (ESO), have therefore promoted stan-

Address for Correspondence: Hatice Yelda Yıldız, Department of Neurology Stroke Center, İstinye University Faculty of Medicine, İstanbul, Türkiye
E-mail: yeldayildiz82@hotmail.com **ORCID ID:** 0009-0002-2452-1874

Received date: 10.08.2026
Revised date: 09.09.2026
Accepted date: 13.09.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

dardized quality indicators and data-collection frameworks to support benchmarking, auditing, and continuous quality improvement across healthcare systems.[3,4]

Within health systems, every component of stroke care—from team composition and equipment standards to treatment protocols—is explicitly defined and periodically updated to align with new scientific developments. Ensuring the reliability of these evolving structures requires integration into quality-assurance frameworks that measure time-sensitive interventions, complication rates, and the implementation of secondary-prevention strategies. These standards cover acute treatment, post-acute care, and rehabilitation, enabling consistent performance monitoring over time.[5]

To address variations among regions and levels of resource availability, the ESO-EAST initiative launched the Registry of Stroke Care Quality (RES-Q) platform in 2015. This web-based data-collection system was designed to analyse healthcare quality indicators, provide hospital- and national-level feedback and benchmarking, and support healthcare systems in monitoring adherence to internationally accepted performance measures.[6] RES-Q therefore embodies a comprehensive and adaptable quality-assessment framework, optimized to accommodate diverse healthcare environments. For such systems to function effectively, certification and auditing mechanisms must rely on clearly defined, measurable performance criteria, assessed at regular intervals by independent evaluators. Infrastructure, clinical protocols, and workflow processes should all be encompassed within these evaluations. Collectively, these interlocking measures constitute a foundation for continuous quality improvement in stroke care worldwide.

Against this background, the {blinded} University Stroke Center integrated selected RES-Q-compatible documentation fields into its Hospital Information System (HIS) to establish a continuous, data-based quality assessment model. This study aims to evaluate the implementation of RES-Q-compatible structured documentation within the HIS, assess changes in documentation completeness for selected quality indicators, and provide insights into developing a sustainable quality-monitoring framework for stroke care in Turkey. The study was not designed as a full RES-Q registry validation project but rather as an assessment of HIS-based documentation quality and RES-Q compatibility.

MATERIALS and METHODS

Study Design

This single-center observational quality-improvement study used a before-and-after design to evaluate the completeness of selected RES-Q-compatible documentation fields before and after integration of structured electronic templates into the HIS. A retrospective pre-integration period (November 2022–October 2023) was compared with a prospective post-integration period (January–December 2024). November–December 2023 represented the implementation and system optimization period and was therefore excluded from comparative analyses.

Study Population

All consecutive stroke admissions recorded within the institutional stroke workflow during the study periods were included. No patient-level exclusion criteria were applied because the objective of the study was to evaluate documentation completeness rather than clinical outcomes. The unit of analysis was the documentation record associated with each stroke admission. No formal sample-size calculation was performed because all available stroke records during the study periods were included.

A total of 2,326 stroke admissions were available during the study periods, comprising 1,565 records from the pre-integration period and 761 records from the post-integration period.

Setting and Organization of Stroke Care

The İstinye University Stroke Center has been structured in accordance with comprehensive stroke center standards, providing the full continuum of care from emergency admission to post-stroke rehabilitation. The patient triage and management workflow is regulated by the institutional protocol entitled "Nörokod – Acute Stroke Diagnosis and Treatment Process," which ensures real-time coordination among vascular neurologists, the ward physician (stroke-floor physician), and the Nörokod nurse.

The stroke-floor physician is a general practitioner with a minimum of one year of experience in emergency medicine and specialized training in neurological examination, the National Institutes of Health Stroke Scale (NIHSS), hyperacute therapies, and stroke-specific imaging standards. This physician is responsible for the initial patient assessment, collection of medical history, initiation of emergency treatment, and documentation of all clinical actions.

The Nörokod nurse assists in recording vital signs, reporting critical values, facilitating intra-hospital transfers, and coordinating the diagnostic and treatment processes with the neurologist and ward physician. During radiological examinations, the nurse evaluates potential complications, measures critical time metrics, and records them in real time.

Data Sources and Measurement

Patient histories and clinical data are recorded using standardized electronic templates aligned with RES-Q quality indicators. Data entry is performed directly at the point of care by trained clinical staff to prevent retrospective input or secondary transcription errors, thereby reinforcing data reliability and enhancing quality assurance within the workflow.

The objective of the project was not to create a parallel RES-Q registry database or duplicate all RES-Q variables within a separate form. Instead, the project focused on embedding RES-Q-compatible structured documentation into routine clinical workflows. Variables already available in institutional systems, including many radiological, laboratory, and timestamp-related parameters, remained accessible through existing HIS, PACS, and laboratory information systems and therefore did not require duplicate manual registration.

The study was structured around two comparative periods separated by an implementation interval:

Pre-integration period – Retrospective evaluation (November 1, 2022–October 30, 2023): Retrospective records were analyzed to determine baseline documentation completeness for predefined RES-Q-compatible clinical indicators.

Implementation interval – System design and optimization (November–December 2023): RES-Q-compatible structured templates were developed, integrated into the HIS workflow, and optimized for routine clinical use. This interval was excluded from comparative analyses.

Post-integration period – Prospective evaluation (January 1, 2024–December 31, 2024): Following system refinement, prospective data collection was conducted to assess changes in documentation completeness after RES-Q-compatible HIS integration.

To minimize information bias, data were entered at the point of care using standardized electronic templates and predefined response options whenever applicable.

Scope of RES-Q Indicators Included in the Analysis

Although the RES-Q framework encompasses a broad spectrum of quality indicators across acute stroke management, imaging, reperfusion therapies, rehabilitation, complications, and secondary prevention, the purpose of the present study was not to evaluate the completeness of the entire RES-Q dataset.

The study specifically focused on clinical documentation domains in which structured recording by healthcare professionals was required and where improvements in documentation quality could be directly influenced by HIS integration. Several RES-Q-related variables, including many radiological findings, imaging timestamps, laboratory measurements, and other digitally generated parameters, were already available through institutional information systems and therefore were not included in the completeness analysis.

Consequently, the present study should be interpreted as an evaluation of RES-Q-compatible documentation quality within the HIS environment rather than a comprehensive assessment of all RES-Q registry variables.

System Integration and Data Handling

The project aimed to embed RES-Q quality indicators within the HIS architecture, allowing structured data capture from electronic templates without the need for a separate parallel RES-Q form. The HIS-based workflow encompassed every stage of care—from patient admission to discharge—ensuring data traceability and real-time quality monitoring.

As emphasized in prior literature, continuously updated patient registries provide key advantages by allowing expansion of sample size through ongoing case inclusion and longitudinal trend analyses across years.[1,5,6]

This study thus evaluates the applicability and effectiveness of RES-Q-based structured templates as fundamental components of the HIS, specifically assessing their capacity to enhance data integrity, standardize documentation, and facilitate non-neurologist stroke-team participation in high-quality registry-based recording aligned with international quality metrics.

Outcome Definition

The primary outcome was documentation completeness, defined as the proportion of eligible records containing a valid documented value for each selected RES-Q-compatible variable. Completeness rates were compared between the pre-integration and post-integration periods. The analysis assessed documentation completeness rather than clinical outcomes, treatment effectiveness, or patient prognosis.

The evaluated variables were selected RES-Q-compatible documentation fields grouped into four predefined domains: Stroke Circumstances, Previous Known History and Comorbidities, Pre-Admission Treatment, and Clinical Measures. For each variable, documentation completeness was assessed as the presence or absence of a valid recorded entry. For pre-admission treatment variables, a "filled" record indicated that medication status was explicitly documented in the corresponding structured field, irrespective of whether the patient was using the medication. Thus, medication-re-

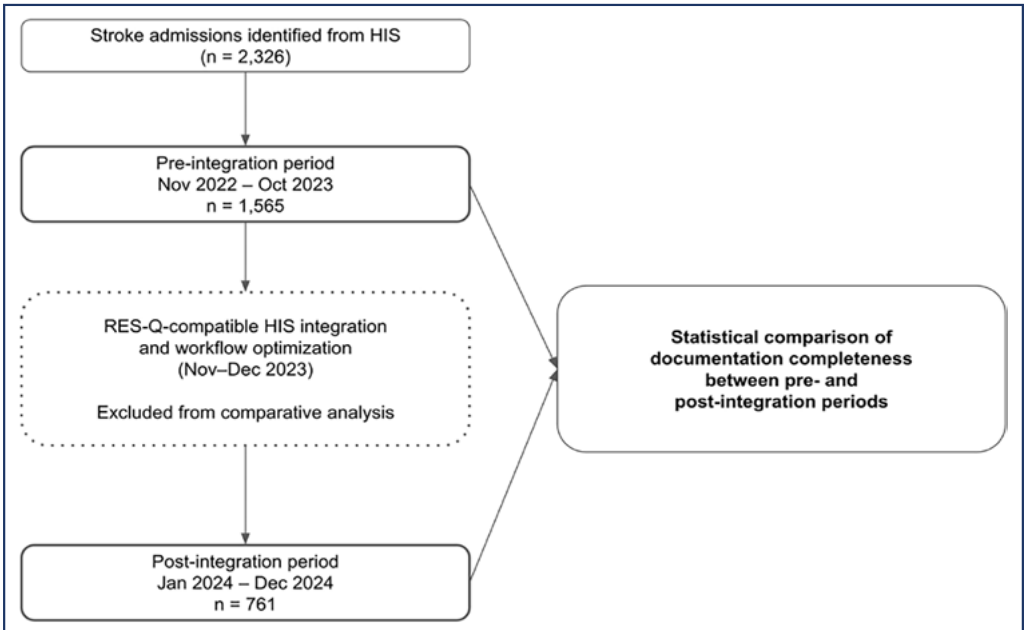


Figure 1. Study flow diagram

Flow diagram showing identification of stroke admissions, implementation of RES-Q-compatible structured documentation within the hospital information system (HIS), and comparative assessment of documentation completeness between the pre-integration (November 2022–October 2023) and post-integration (January–December 2024) periods. The implementation and system optimization phase (November–December 2023) was excluded from comparative analyses

lated completeness rates reflect the availability of documented medication-status information rather than the prevalence of medication use.

Missing Data

Missing values were not imputed. For each variable, completeness was calculated using the proportion of records containing a documented value relative to all eligible stroke records within the corresponding study period.

Statistical Analyses and Tools

Clinical data profiling was conducted using Wanalyzer v3.0 (WisdomEra Corp., Istanbul, Turkey), a platform integrated into the HIS (https://wisdomera.io). The application incorporates natural language processing (NLP) and expression management systems to extract, structure, and classify clinical information from unstructured electronic medical records. These features enabled the mapping of clinical narratives to standardized RES-Q-compatible quality indicators.

Statistical analyses were performed using Wistats v3.0 (WisdomEra Corp., Istanbul, Turkey), which operates on the Python programming language and utilizes SciPy v1.2.3, scikit-learn v0.24.0, and statsmodels v0.9.0 (https://wistats.wisdomera.io). Age is presented as mean±standard deviation and was compared between study periods using the independent samples t-test. Categorical variables are presented as frequencies and percentages. Comparisons of documentation completeness between the pre-integration and post-integration periods were performed using the chi-square test or Fisher's exact test, when appropriate.

The primary analysis evaluated the proportion of completed documentation records for each selected RES-Q-compatible variable before and after HIS integration. Completeness rates were calculated as the percentage of records containing a valid entry among all eligible stroke admissions within the corresponding study period.

All statistical tests were two-sided, and a p value<0.05 was considered statistically significant.

RESULTS

Study Flow and Included Records

The study flow is presented in Figure 1. During the study period, a total of 2,326 stroke cases were documented in the RES-Q-compatible HIS. Of these, 1,565 records belonged to the retrospective pre-integration period (November 2022–October 2023), and

Table 1. Comparison of documentation completeness for stroke circumstances before and after res-q-compatible his integration

Variables	Before mean / %	After mean / %	p
Number of cases	1565 (100)	761 (100)	
Gender			0.059
Female	679 (43.4)	357 (46.9)	
Male	886 (56.6)	404 (53.1)	
Age	71.1	70.6	0.465
Stroke while already hospitalized			<0.001
Filled	1031 (65.9)	575 (75.6)	
Not-Filled	534 (34.1)	186 (24.4)	
Wake up stroke			<0.001
Filled	729 (46.6)	649 (85.3)	
Not-Filled	836 (53.4)	112 (14.7)	
Time of admission (hospital door)			
Filled	1565 (100.0)	761 (100.0)	
Not-Filled	0 (0.0)	0 (0.0)	
Date of onset of stroke symptoms			<0.001
Filled	900 (57.5)	680 (89.4)	
Not-Filled	665 (42.5)	81 (10.6)	
First assessment location			
Filled	1565 (100.0)	761 (100.0)	
Not-Filled	0 (0.0)	0 (0.0)	
Arrival Source			<0.001
Filled	1199 (76.6)	702 (92.2)	
Not-Filled	366 (23.4)	59 (7.8)	
Name of first hospita of admission			<0.001
Filled	1176 (75.1)	647 (85.0)	
Not-Filled	389 (24.9)	114 (15.0)	
Patient hospitalized in Day 1			
Filled	1565 (100.0)	761 (100.0)	
Not-Filled	0 (0.0)	0 (0.0)	
Admitted under department			
Filled	1565 (100.0)	761 (100.0)	
Not-Filled	0 (0.0)	0 (0.0)	

Values are presented as n (%) or mean values. Documentation completeness represents the proportion of records containing a valid documented value for each variable. Comparisons were performed between the pre-integration period (November 2022–October 2023) and the post-integration period (January–December 2024). The implementation interval (November–December 2023) was excluded from analysis. Categorical variables were compared using the Chi-square test or Fisher's exact test, and continuous variables using the Independent Samples t-test. HIS: Hospital information system; RES-Q: Registry of stroke care quality

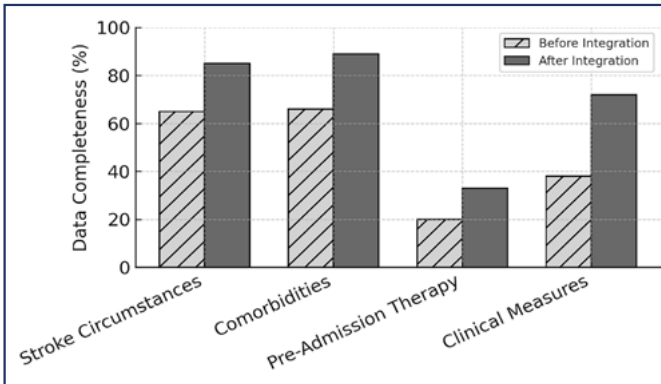


Figure 2. Overall trend in data completeness (before vs after integration)

This figure illustrates the improvement in data completeness across four analyzed RES-Q-compatible documentation domains—Stroke Circumstances, Comorbidities, Pre-Admission Therapy, and Clinical Measures—before and after integration of selected RES-Q-compatible templates into the hospital information system (HIS)

761 records were captured during the prospective post-integration period (January–December 2024). The implementation and system optimization interval (November–December 2023) was excluded from comparative analyses. All available entries from both periods were included in the analysis without exclusions. Because the study focused on documentation completeness rather than patient follow-up or clinical outcomes, no eligibility screening or attrition stages were applicable.

Stroke Circumstances

In the Stroke Circumstances domain, the proportion of completed records increased significantly for nearly all analyzed parameters. Documentation of stroke occurring during hospitalization improved from 65.9% to 75.6%, wake-up stroke from 46.6% to 85.3%, and symptom-onset time from 57.5% to 89.4% (all p<0.001; Table 1). Similarly, completeness for arrival source increased from 76.6% to 92.2%, and first hospital of admission from 75.1% to 85.0% (both p<0.001; Table 1). Overall domain-level improvement is summarized in Figure 2.

Previous Medical History and Comorbidities

Completeness of previous medical history and comorbidity documentation improved substantially after RES-Q-compatible HIS integration. Filled records increased for hypertension from 78.8% to 93.6%, coronary artery disease or myocardial infarction from 61.1% to 92.6%, diabetes mellitus from 67.0% to 91.9%, congestive heart failure from 56.9% to 83.7%, and atrial fibrillation or flutter from 64.8% to 86.6%

Table 2. Comparison of documentation completeness for previous medical history and comorbidities before and After RES-Q-compatible HIS integration

Variables	Before n (%)	After n (%)	p
Number of cases	1565 (100)	761 (100)	
Hypertension			<0.001
Filled	1234 (78.8)	712 (93.6)	
Not-filled	331 (21.2)	49 (6.4)	
Coronary artery disease			<0.001
Filled	956 (61.1)	705 (92.6)	
Not-filled	609 (38.9)	56 (7.4)	
Diabetes			<0.001
Filled	1049 (67.0)	699 (91.9)	
Not-filled	516 (33.0)	62 (8.1)	
Congestive heart failure			<0.001
Filled	891 (56.9)	637 (83.7)	
Not-filled	674 (43.1)	124 (16.3)	
Hormonal contraception			
Filled	0 (0.0)	0 (0.0)	
Not-filled	1565 (100.0)	761 (100.0)	
Smoking			<0.001
Filled	475 (30.4)	20 (2.6)	
Not-filled	1090 (69.6)	741 (97.4)	
HIV			<0.001
Filled	745 (47.6)	230 (30.2)	
Not-filled	820 (52.4)	531 (69.8)	
Previous hemorrhagic stroke			<0.001
Filled	862 (55.1)	236 (31.0)	
Not-filled	703 (44.9)	525 (69.0)	
Atrial fibrillation/flutter			<0.001
Filled	1014 (64.8)	659 (86.6)	
Not-filled	551 (35.2)	102 (13.4)	

Data are presented as n (%). Documentation completeness represents the proportion of records containing a valid documented value for each comorbidity variable. Comparisons were performed between the pre-integration period (November 2022–October 2023) and the post-integration period (January–December 2024). The implementation interval (November–December 2023) was excluded from analysis. P values were calculated using the Chi-square test or Fisher's exact test, as appropriate. RES-: Registry of stroke care quality; HIS: Hospital information system; MI: Myocardial infarction; HIV: Human immunodeficiency virus

(all p<0.001; Table 2). In contrast, completeness declined for smoking status from 30.4% to 2.6%, HIV history from 47.6% to 30.2%, and previous hemorrhagic stroke from 55.1% to 31.0% (all p<0.001; Table 2). Parameter-specific changes are illustrated in Figure 3.

Pre-Admission Treatment

Documentation completeness also improved in the Pre-Admission Treatment domain. Recording of antihypertensive use increased from 25.5% to 38.6%, antidiabetic medication from 11.7% to 19.7%, and statin therapy from 4.9% to 9.5% (all p<0.001; Table 3). Completeness for apixaban increased from 2.2% to 3.7%, edoxaban from 0.6% to 1.4%, and cilostazol from 0.1% to 0.5% (all p<0.05; Table 3).

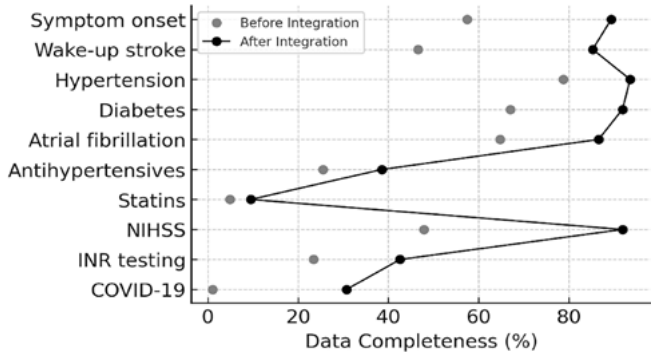


Figure 3. Parameter-specific changes in data completeness

Parameter-level trends in completeness are shown for representative RES-Q variables. Gray circles indicate values before, and black circles indicate values after integration. Consistent upward shifts demonstrate improved documentation for critical measures such as symptom-onset time (57.5 %)

Clinical Measures

The Clinical Measures domain demonstrated the most pronounced gains. Documentation of NIHSS on admission increased from 47.9% to 91.9%, INR testing from 23.4% to 42.6%, and COVID-19 status from 1.0% to 30.7% (all p<0.001; Table 4). In contrast, documentation completeness for pre-admission modified Rankin Scale (mRS) declined from 5.4% to 0.4% (p<0.001; Table 4). The distribution of completeness improvements across analyzed parameters is shown in Figure 4.

Overall Documentation Completeness

Overall, post-integration analyses showed a statistically significant enhancement in documentation completeness across nearly all analyzed RES-Q-compatible documentation domains. These findings support the effectiveness of RES-Q-compatible structured electronic templates for improving documentation quality and data reliability within the HIS workflow.

DISCUSSION

This study demonstrates that embedding selected RES-Q-compatible documentation fields into the institutional workflow of the Stroke Center was associated with measurable improvements in data completeness and documentation quality across multiple domains of stroke care. Following the

Table 3. Comparison of documentation completeness for pre-admission treatment before and after RES-Q-compatible HIS integration

Variables	Before n (%)	After n (%)	p	Variables	Before n (%)	After n (%)	p
Number of cases	1565 (100)	761 (100)		Dipyridamol			
Anti-hypertensives			<0.001	Filled	0 (0.0)	0 (0.0)	
Filled	399 (25.5)	294 (38.6)		Not-filled	1565 (100.0)	761 (100.0)	
Not-filled	1166 (74.5)	467 (61.4)		Warfarin			0.408
Anti-diabetics			<0.001	Filled	37 (2.4)	16 (2.1)	
Filled	183 (11.7)	150 (19.7)		Not-filled	1528 (97.6)	745 (97.9)	
Not-filled	1382 (88.3)	611 (80.3)		LMWH/Heparin			0.223
Aspirin (ASA)			0.068	Filled	17 (1.1)	5 (0.7)	
Filled	439 (28.1)	237 (31.1)		Not-filled	1548 (98.9)	756 (99.3)	
Not-filled	1126 (71.9)	524 (68.9)		Dabigatran			0.415
Cilostazol			0.024	Filled	4 (0.3)	3 (0.4)	
Filled	1 (0.1)	4 (0.5)		Not-filled	1561 (99.7)	758 (99.6)	
Not-filled	1564 (99.9)	757 (99.5)		Rivaroxaban			0.055
Clopidogrel			0.070	Filled	53 (3.4)	37 (4.9)	
Filled	98 (6.3)	61 (8.0)		Not-filled	1512 (96.6)	724 (95.1)	
Not-filled	1467 (93.7)	700 (92.0)		Apixaban			0.032
Ticagrelor				Filled	35 (2.2)	28 (3.7)	
Filled	0 (0.0)	0 (0.0)		Not-filled	1530 (97.8)	733 (96.3)	
Not-filled	1565 (100.0)	761 (100.0)		Edoxaban			0.032
Ticlopidine				Filled	9 (0.6)	11 (1.4)	
Filled	0 (0.0)	0 (0.0)		Not-filled	1556 (99.4)	750 (98.6)	
Not-filled	1565 (100.0)	761 (100.0)		Statin			<0.001
Prasugrel				Filled	76 (4.9)	72 (9.5)	
Filled	0 (0.0)	0 (0.0)		Not-filled	1489 (95.1)	689 (90.5)	
Not-filled	1565 (100.0)	761 (100.0)					

Data are presented as n (%). Documentation completeness represents the proportion of records containing a valid documented value for each pre-admission medication variable. Comparisons were performed between the pre-integration period (November 2022–October 2023) and the post-integration period (January–December 2024). The implementation interval (November–December 2023) was excluded from analysis. P values were calculated using the Chi-square test or Fisher's exact test, as appropriate. RES-Q: Registry of stroke care quality; HIS: Hospital information system; ASA: Acetylsalicylic acid; LMWH: Low-molecular-weight heparin

implementation of structured electronic templates and direct point-of-care entry, we observed substantial gains in domains such as stroke circumstances (e.g., symptom-onset time, arrival source), comorbidities and pre-admission treatments, and key clinical measures (e.g., NIHSS on admission). These findings align with prior evidence indicating that well-designed registries and focused quality systems can enhance both data fidelity and, by extension, care performance.[7,8]

Importantly, our objective was not to replicate the complete RES-Q registry structure within a separate data-entry environment. Instead, we sought to improve the completeness of clinical information that requires active documentation by the stroke team. Many radiological, laboratory, and system-generated quality indicators already existed within the institutional digital infrastructure and therefore did not require additional manual registration. This distinction is important when interpreting the scope of the reported completeness analyses.

These improvements in documentation—for example, increasing the recording of NIHSS on admission from 47.9% to 91.9%—suggest that embedding process measures into routine care can lead to meaningful workflow change. This is consistent with the notion that adherence to performance indicators is linked to better patient outcomes.[9] In other words, enhanced data completeness is not simply a technical achievement but may reflect deeper changes in care culture, from "optional documentation" to "care-system-integrated documentation."

The results also reflect broader insights from the epidemiology of stroke and the evolving need for quality improvement. Globally, despite declining age-standardized rates in some settings, the absolute burden of stroke continues to rise.[10,11] Thus, establishing robust quality-monitoring systems such as RES-Q is a critical response. Additionally, as De Cola et al.[12] highlight, quality indicators need to span not only acute care but also rehabilitation and follow-up phases to ensure

Table 4. Comparison of documentation completeness for clinical measures before and after RES-Q-compatible HIS integration

Variables	Before n (%)	After n (%)	p
Number of cases	1565 (100)	761 (100)	
NIHSS on admission			<0.001
Filled	750 (47.9)	699 (91.9)	
Not-filled	815 (52.1)	62 (8.1)	
Glasgow coma scale			<0.001
Filled	121 (7.7)	701 (92.1)	
Not-filled	1444 (92.3)	60 (7.9)	
mRS			<0.001
Filled	85 (5.4)	3 (0.4)	
Not-filled	1480 (94.6)	758 (99.6)	
COVID-19 status			<0.001
Filled	16 (1.0)	234 (30.7)	
Not-filled	1549 (99.0)	527 (69.3)	
INR testing			<0.001
Filled	366 (23.4)	324 (42.6)	
Not-filled	1199 (76.6)	437 (57.4)	

Data are presented as n (%). Documentation completeness represents the proportion of records containing a valid documented value for each clinical measure. Comparisons were performed between the pre-integration period (November 2022–October 2023) and the post-integration period (January–December 2024). The implementation interval (November–December 2023) was excluded from analysis. P values were calculated using the Chi-square test or Fisher's exact test, as appropriate. RES-Q: Registry of stroke care quality; NIHSS: National institutes of health stroke scale; INR: International normalized ratio; mRS: Modified rankin scale; COVID-19: Coronavirus disease 2019

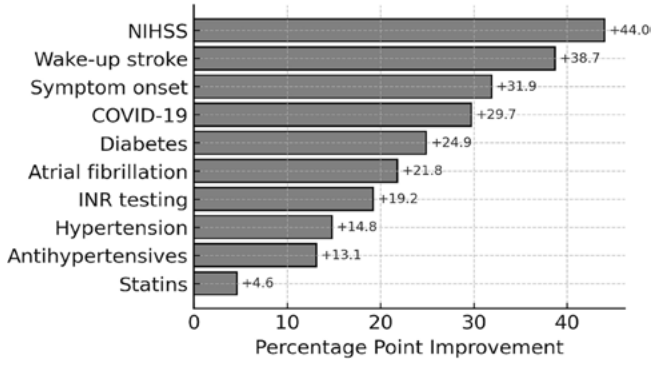


Figure 4. Completeness improvement distribution (after–before)

This horizontal bar graph displays the percentage-point change (Δ %) in data completeness for all analyzed parameters, ranked from greatest to smallest improvement. The most pronounced gains were observed in NIHSS on admission (+44 pp), symptom-onset documentation (+32 pp), wake-up stroke (+39 pp), and diabetes mellitus (+25 pp). Gray gradient bars highlight the magnitude of improvement for each metric, demonstrating the overall enhancement in data recording performance following RES-Q-compatible HIS integration

that the full stroke-care continuum is captured. Our HIS template structure was designed to reflect this broad continuum; however, the present analysis focused on selected clinical documentation domains.

Nevertheless, some limitations and key challenges remain. In this study, certain parameters, including smoking status, HIV history, previous hemorrhagic stroke, and pre-admission mRS, showed declining completeness after implementation. This suggests that even when systems are optimized, certain data fields remain vulnerable—likely those that are not directly tied to immediate acute-care workflow or whose documentation responsibilities are less clearly owned by the stroke team. This observation mirrors the findings of Purvis et al.[13] that many process indicators, especially those related to early rehabilitation or non-immediate care phases, are underused or misaligned with guideline-derived sets. The decline in mRS completeness, for example, may reflect difficulties in capturing functional outcomes in real time within busy stroke units.

Moreover, while data completeness is a prerequisite for outcome monitoring and benchmarking, it is not in itself a guarantee of care quality improvement or better patient outcomes. For example, documentation improvements do not directly equate to faster door-to-needle times, shorter lengths of stay, or lower recurrence rates; these downstream endpoints require separate measurement and may lag behind documentation changes. In this respect, our study should be interpreted as an intermediate outcome (i.e., documentation quality) rather than a patient-outcome study. Future work should link improved documentation to actual clinical outcomes and process times (e.g., reperfusion delays, complication rates) to complete the "quality-improvement loop."

From a systems perspective, our experience reinforces three major enablers for success: (1) integration of selected RES-Q-compatible templates into the HIS such that data entry is seamless, (2) training and empowerment of multidisciplinary team members (including non-neurologist physicians, nurses, and information-system staff) to adopt the new workflow, and (3) feedback loops and periodic auditing to sustain performance. Our institution's before-and-after approach, with a defined implementation interval between retrospective baseline assessment and prospective monitoring, may provide a replicable model for other centers seeking to operationalize registry-based quality improvement.

Another important consideration is that the study evaluated selected RES-Q-compatible documentation domains rather

than the entire spectrum of RES-Q indicators. This approach was intentional because numerous imaging- and laboratory-related parameters were already routinely captured by existing hospital systems. Therefore, the findings should be interpreted as reflecting improvements in structured clinical documentation rather than completeness of every variable included in the RES-Q framework.

Despite these strengths, several limitations should be acknowledged. Although the integration of selected RES-Q-compatible documentation fields into the HIS significantly improved documentation completeness, this study has several limitations. First, data quality in the pre-integration period depended on retrospective documentation, which may introduce information bias and limit the accuracy of baseline completeness estimates. Second, while the structured templates enhanced standardization, variations in clinician engagement and workflow intensity could still influence data entry reliability, leading to differential completeness across indicators. Third, the study evaluated documentation quality rather than clinical outcomes; thus, improvements in completeness cannot be directly linked to better patient care or functional recovery. Additionally, the findings reflect the experience of a single stroke center, which may limit generalizability to other institutions with different staffing models or HIS infrastructures. Finally, the study did not assess the long-term sustainability of improvements, and further follow-up is required to determine whether documentation gains persist over time.

In the context of the Turkish healthcare environment, where stroke center certification and systematic registries are less uniformly established than in some high-income countries, our findings may have broader implications. Implementing the RES-Q template structure may help bridge regional disparities, support benchmarking, and foster continuous data-driven improvement in stroke care. Given the rising stroke burden projected globally and in middle-income settings, such infrastructures are timely and necessary.[14]

CONCLUSION

In conclusion, while documentation completeness is one component of care quality, the substantial gains observed in our center following the integration of selected RES-Q-compatible templates into the daily workflow suggest that systematically applied quality-registry frameworks can materially improve data capture. Sustaining gains will require ongoing training, performance feedback, integration across the full stroke pathway (including rehabilitation), and measurement of whether improved documentation translates into improved clinical outcomes.

Disclosures

Ethics Committee Approval: The study was approved by the Istinye University Human Research Ethics Committee (No: 24-272, Date: 22/11/2024).

Informed Consent: Informed consent was obtained from all participants.

Conflict of Interest Statement: None of the authors has any potential financial conflict of interest related to this manuscript.

Funding: This research received no external funding.

Use of AI for Writing Assistance: No artificial intelligence-assisted technologies were used in the preparation, analysis, or writing of this manuscript.

Author Contributions: Concept – H.Y.Y., Y.K.; Design – H.Y.Y., N.O., N.Ü., U.E., F.K.Z.; Supervision – H.Y.Y., Y.K., Y.B., Ş.K., T.Ö.; Data collection and/or processing – N.O., N.Ü., T.Ö., Y.B., Ş.K., U.E., F.K.Z.; Data analysis and/or interpretation – H.Y.Y.; Literature search – H.Y.Y., Y.K.; Writing – H.Y.Y.; Critical review – H.Y.Y., Y.K., N.O., N.Ü., U.E., F.K.Z., Y.B., Ş.K., T.Ö.

Acknowledgments: We would like to thank the Artificial Intelligence and Research Center of Istinye University (https://yzaum.istinye.edu.tr/, accessed on 06 August 2026) for providing technical support during manuscript preparation, including assistance with data validation processes and pre-submission similarity evaluation. The authors also acknowledge the Ditako Data Analytics Team (https://ditako.com, accessed on 06 August 2026) for their valuable contributions to the statistical analyses and methodological support provided for this study.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Feigin VL, Brainin M, Norrving B, Martins S, Sacco RL, Hacke W, et al. World Stroke Organization (WSO): Global Stroke Fact Sheet 2022. Int J Stroke Off J Int Stroke Soc 2022;17:18–29. [CrossRef]
2. Feigin VL, Owolabi MO; World Stroke Organization-Lancet Neurology Commission Stroke Collaboration Group. Pragmatic solutions to reduce the global burden of stroke: a World Stroke Organization-Lancet Neurology Commission. Lancet Neurol 2023;22:1160–206. [CrossRef]
3. Ringelstein EB, Chamorro A, Kaste M, Langhorne P, Leys D, Lyrer P, et al; ESO Stroke Unit Certification Committee. European Stroke Organisation recommendations to establish a stroke unit and stroke center. Stroke 2013;44:828–40. [CrossRef]
4. Stroke Unit Trialists' Collaboration. Organised inpatient (stroke unit) care for stroke. Cochrane Database Syst Rev 2013;2013:CD000197.
5. Gropen TI, Gagliano PJ, Blake CA, Sacco RL, Kwiatkowski T, Richmond NJ, et al. Quality improvement in acute stroke: the New York State Stroke Center Designation Project. Neurology 2006;67:88–93. [CrossRef]
6. Mikulík R, Caso V, Bornstein NM, Svobodová V, Pezzella FR, Grecu A, et al. Enhancing and accelerating stroke treatment in Eastern European region: Methods and achievement of the ESO EAST program. Eur Stroke J 2020;5:204–12. [CrossRef]

7. Stewart VT. Use of a prototype acute stroke registry to improve care: profile of receptive stroke programs. Am J Prev Med 2006;31(6 Suppl 2):S217–223. [CrossRef]
8. Lanzola G, Parimbelli E, Micieli G, Cavallini A, Quaglini S. Data quality and completeness in a web stroke registry as the basis for data and process mining. J Healthc Eng 2014;5:163–84. [CrossRef]
9. Urimubenshi G, Langhorne P, Cadilhac DA, Kagwiza JN, Wu O. Association between patient outcomes and key performance indicators of stroke care quality: A systematic review and meta-analysis. Eur Stroke J 2017;2:287–307. [CrossRef]
10. Katan M, Luft A. Global burden of stroke. Semin Neurol 2018;38:208–11. [CrossRef]
11. GBD 2019 Stroke Collaborators. Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021;20:795–820. [CrossRef]
12. De Cola MC, Ielo A, Lo Buono V, Quartarone A, Calabrò RS. Toward social-health integration in Sicily: description of the first hub and spoke model to improve the diagnostic therapeutic care paths for neurorehabilitation. Front Public Health 2023;11:1141581. [CrossRef]
13. Purvis T, Cadilhac D, Donnan G, Bernhardt J. Systematic review of process indicators: including early rehabilitation interventions used to measure quality of acute stroke care. Int J Stroke Off J Int Stroke Soc 2009;4:72–80. [CrossRef]
14. Feigin VL, Brainin M, Norrving B, Martins SO, Pandian J, Lindsay P, et al. World Stroke Organization: Global Stroke Fact Sheet 2025. Int J Stroke 2025;20:132–44. [CrossRef]

DOI: 10.14744/cm.2026.44712
Compreh Med 2026;18(Suppl. 1):94-102

RESEARCH ARTICLE

Association of Lithium Treatment with Biological Rhythms in Euthymic Patients with Bipolar Disorder

Ayşe Kurtulmuş, Fatma Büşra Parlakkaya, Halenur Şamandar, Pınar Ertunç, Zülal Çelik

Department of Psychiatry, İstanbul Medeniyet University Faculty of Medicine, İstanbul, Türkiye

ABSTRACT

Objective: Circadian rhythm disturbances are increasingly recognized as a core feature of bipolar disorder. Although lithium is a gold standard of maintenance treatment, clinical evidence regarding its association with sleep and biological rhythms remains limited. This study compared sleep quality, chronotype, and biological rhythms between euthymic patients with bipolar disorder receiving lithium and those under other treatments.

Materials and Methods: In this cross-sectional study, 83 euthymic outpatients with bipolar disorder (32 receiving lithium and 51 not receiving lithium) completed the Pittsburgh Sleep Quality Index (PSQI), Morningness–Eveningness Questionnaire (MEQ), and Biological Rhythms Interview of Assessment in Neuropsychiatry (BRIAN). Multiple linear regression analyses examined whether lithium treatment was independently associated with biological rhythm measures after adjustment for age, sex, and subjective sleep quality. Associations between serum lithium concentrations and sleep/circadian measures were examined within the lithium group. Sensitivity analyses were performed using a therapeutic serum lithium concentration threshold (≥0.6 mmol/L).

Results: After adjustment for age, sex, and subjective sleep quality, lithium treatment was independently associated with lower BRIAN total scores (p=0.048) but not with MEQ scores. Serum lithium concentrations were not linearly associated with PSQI, MEQ, or BRIAN scores. In sensitivity analyses, therapeutic lithium exposure was associated with lower BRIAN total scores (p=0.011) and lower sleep, social, and eating domain scores (p=0.039, p=0.029, and p=0.036, respectively).

Conclusion: Lithium treatment was independently associated with lower biological rhythm disturbance, but not with subjective sleep quality or chronotype, in euthymic patients with bipolar disorder. These findings are consistent with the hypothesis that lithium treatment may be associated with greater biological rhythm stability beyond differences in subjective sleep quality; however, longitudinal studies are needed to clarify the temporal nature of this relationship.

Keywords: Bipolar disorder, circadian rhythms, lithium, sleep

How to cite this article: Kurtulmuş A, Parlakkaya FB, Şamandar H, Ertunç P, Çelik Z. Association of Lithium Treatment with Biological Rhythms in Euthymic Patients with Bipolar Disorder . Compreh Med 2026;18(Suppl. 1):94-102.

INTRODUCTION

Bipolar disorder (BD) is a chronic, recurrent psychiatric disorder affecting approximately 1% of the population worldwide and is among the leading causes of disability in young adults.[1] It is characterized by recurrent episodes of mania, hypomania, depression, or mixed states that substantially impair psychosocial functioning and quality of life. Although acute affective episodes define the clinical course of the disorder, many patients continue to experience residual symptoms, cognitive difficulties, and functional impairment during euthymic periods, highlighting that bipolar disorder is not merely an episodic illness but a chronic condition with a persistent clinical burden.[2–4]

Sleep disturbance is one of the core clinical features of bipolar disorder and represents a hallmark symptom across all phases of the illness. Reduced need for sleep is a diagnostic feature of mania, whereas insomnia and hypersomnia are common manifestations of bipolar depression. Moreover, sleep disturbance has been demonstrated to be prevalent during euthymia, even in patients on maintenance medication, and has been associated with poorer functioning, lower quality of life, and a greater risk of mood episode recurrence.[5,6] These findings suggest that sleep abnormalities may represent a trait-like feature of bipolar disorder rather than simply a state-dependent manifestation of acute mood episodes.

Address for Correspondence: Ayse Kurtulmuş, Department of Psychiatry, İstanbul Medeniyet University Faculty of Medicine, İstanbul, Türkiye
E-mail: a_kurtulmus@hotmail.com **ORCID ID:** 0000-0001-8420-4761

Received date: 11.07.2026
Revised date: 10.09.2026
Accepted date: 15.09.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Sleep is regulated by the interaction of two complementary physiological processes: the homeostatic sleep drive and the circadian timing system. The homeostatic process is the accumulation of sleep pressure during wakefulness, and the circadian system synchronizes sleep–wake timing to the external light–dark cycle via the central biological clock in the suprachiasmatic nucleus.[7] The circadian system regulates the sleep–wake cycle and coordinates numerous physiological functions, such as endocrine secretion, metabolism, immune function, autonomic activity, cognitive performance, and emotional regulation.[8,9] Therefore, circadian disruptions have widespread effects extending beyond sleep itself.

Emerging evidence indicates that disturbances in circadian rhythm are a core feature of bipolar disorder. Bipolar disorder patients display abnormalities in sleep–wake timing, patterns of daily activity, melatonin secretion, and other circadian markers, many of which persist in euthymia.[10–13] In addition, genetic studies have consistently implicated several circadian clock genes, including CLOCK, ARNTL (BMAL1), PER, and CRY genes, in susceptibility to bipolar disorder, further supporting the involvement of circadian dysregulation in its pathophysiology.[14]

Lithium remains the gold-standard maintenance treatment for bipolar disorder and is one of the few pharmacological agents shown to be effective in preventing both manic and depressive recurrences while reducing suicide risk.[15] Despite decades of clinical use, its mechanisms of action are not yet fully understood and appear to differ from those of other mood stabilizers and psychotropic medications. In experimental studies, lithium has been shown to directly interact with the molecular circadian clock through inhibition of glycogen synthase kinase-3β (GSK-3β), causing changes in clock gene expression and stabilization of circadian rhythmicity.[16] These findings have led to the hypothesis that modulation of the circadian system may be responsible for the therapeutic effects of lithium. However, evidence from clinical studies investigating whether lithium treatment is associated with improved biological rhythm regulation, especially independently of its effects on sleep, remains limited. Distinguishing the effects of lithium on biological rhythms from its effects on sleep may provide further insight into the mechanisms underlying its mood-stabilizing properties.

Therefore, the present study aimed to compare sleep quality, chronotype, and biological rhythm disturbances between euthymic patients with bipolar disorder receiving lithium and those not receiving lithium. In addition, we examined whether lithium treatment was independently associated with biological rhythm disturbance after controlling for subjective sleep quality. Finally, among patients receiving lithium, we explored the relationship between serum lithium concentrations and sleep and circadian rhythm measures and performed a sensitivity analysis based on therapeutic serum lithium concentrations.

MATERIALS and METHODS

Study Design and Participants

This cross-sectional study included euthymic patients with bipolar disorder who were consecutively recruited from the outpatient psychiatry clinic of Istanbul Medeniyet University between October 2025 and May 2026. The study protocol was approved by the Istanbul Medeniyet University Göztepe Prof. Dr. Süleyman Yalçın City Hospital Clinical Research Ethics Committee (Approval No: 2025/0208; Date: 09.10.2025). All participants provided written informed consent before enrollment. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki and its subsequent revisions. Eligible participants were adults aged 18–65 years with a diagnosis of bipolar disorder who were receiving outpatient psychiatric treatment and were euthymic at the time of assessment. Additional inclusion criteria were the ability to read and write and the provision of written informed consent. Participants were excluded if they had: 1) a diagnosis of a schizophrenia spectrum disorder, 2) alcohol or substance use disorder, 3) intellectual disability, 4) a neurological disorder affecting cognitive functioning (e.g., dementia), 5) a shift-work schedule, 6) transmeridian travel within 2 weeks, or 7) unstable medical illnesses that could substantially affect sleep or circadian rhythm.

Diagnoses were confirmed by experienced psychiatrists according to the Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5). Mood symptom severity was assessed using the Young Mania Rating Scale (YMRS) and the 17-item Hamilton Depression Rating Scale (HAM-D-17). Euthymia was defined as scores ≤7 on both the 17-item Hamilton Depression Rating Scale (HAM-D-17) and the Young Mania Rating Scale (YMRS). One participant in the lithium-treated group was excluded because of subsyndromal depressive symptoms. Consequently, all participants met the study definition of euthymia at the time of assessment.

Participants were classified into two groups according to whether they were receiving lithium treatment as part of their routine maintenance pharmacotherapy. Patients receiving lithium, either as monotherapy or in combination with other psychotropic medications, were assigned to the lithium group.

Patients not receiving lithium, irrespective of their other psychotropic medications, comprised the non-lithium group.

A power calculation based on the primary multiple regression model indicated that a minimum sample size of 81 participants would be required to detect a small-to-moderate effect (Cohen's f^2=0.10) with 80% power at a two-sided α level of 0.05, assuming one tested predictor and four predictors in the full model. The calculation was based on lithium treatment as the predictor of interest, with age, sex, and subjective sleep quality included as covariates.

Clinical Assessment

Sociodemographic and clinical information, including age, sex, educational level, marital status, smoking status, body mass index (BMI), medical comorbidities, and current psychotropic medications, was recorded using a structured clinical form.

Residual depressive and manic symptom severity were assessed using the 17-item Hamilton Depression Rating Scale (HAM-D) and the Young Mania Rating Scale (YMRS), respectively.[17,18]

Serum lithium concentrations were obtained from routine therapeutic drug monitoring performed during clinical follow-up. For patients receiving lithium treatment, the serum lithium concentration closest to the study assessment was recorded. For participants without a serum lithium measurement within the preceding three months, a routine therapeutic drug monitoring assessment was performed as part of standard clinical care. To explore whether therapeutic lithium exposure was associated with sleep and biological rhythm measures, an additional sensitivity analysis was performed using a serum lithium concentration threshold of 0.6 mmol/L, which was considered the lower limit of the therapeutic range.

Assessment of Sleep and Biological Rhythms

Pittsburgh Sleep Quality Index (PSQI)

Subjective sleep quality during the previous month was assessed using the Pittsburgh Sleep Quality Index (PSQI). The 19-item PSQI consists of seven component scores evaluating subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleep medication, and daytime dysfunction.[19] Total scores range from 0 to 21, and higher scores indicate poorer subjective sleep quality. The Turkish version of the PSQI has demonstrated good validity and reliability, with a Cronbach's α value of 0.80.[20]

Morningness–Eveningness Questionnaire (MEQ)

Chronotype was assessed using the Morningness–Eveningness Questionnaire (MEQ). Higher scores indicate a greater preference for morningness, whereas lower scores indicate greater eveningness.[21] In general, scores of 59–86 indicate morning type, 42–58 intermediate type, and 16–41 evening type. The Turkish validity and reliability study was conducted by Agargun et al.,[22] with a Cronbach's α value of 0.81. Because of the moderate sample size of the study, MEQ scores were analyzed as a continuous variable rather than categorizing participants into chronotype groups.

Biological Rhythms Interview of Assessment in Neuropsychiatry (BRIAN)

Biological rhythm disturbances were evaluated using the Biological Rhythms Interview of Assessment in Neuropsychiatry (BRIAN). The BRIAN assesses disturbances across four domains, including sleep, activity, social rhythms, and eating patterns. Three additional items assess individual chronotype preference.[23] The total scale score is calculated by summing the subscale scores. Higher scores indicate greater disruption of biological rhythms. The Turkish validity and reliability study was conducted by Aydemir et al.,[24] with a Cronbach's α value of 0.90.

Statistical Analysis

Statistical analysis were performed using IBM SPSS Statistics version 30.0.0 (IBM Corp., Armonk, NY, USA). Data are presented as mean±standard deviation (SD) for continuous variables and frequencies (percentages) for categorical variables. Normality was assessed using visual inspection of histograms and Q-Q plots together with the Shapiro-Wilk test. Baseline sociodemographic and clinical characteristics were compared between the lithium and non-lithium groups using independent-samples t tests or Welch's t tests for continuous variables, as appropriate, and χ^2 tests for categorical variables. Cohen's d was calculated to estimate effect sizes for continuous variables. Sleep quality (PSQI), chronotype (MEQ), and biological rhythm (BRIAN) scores were initially compared between groups using unadjusted group comparisons. Subsequently, separate multiple linear regression analyses were conducted to determine whether lithium treatment was independently associated with MEQ and BRIAN scores after adjustment for prespecified covariates such as age, sex, and PSQI.

Subjective sleep quality was included as a covariate because the primary aim was to determine whether lithium treatment was associated with biological rhythms beyond differences in subjective sleep quality. Multicollinearity among predictors was assessed using tolerance and variance inflation factor (VIF) values. Among patients receiving lithium treatment,

Table 1. Sociodemographic and clinical characteristics of participants

	Lithium (n=32)	Non-lithium (n=51)	p
Sociodemographic characteristics			
Age, years	36.75±13.26	42.50±11.01	0.036
Female sex, n (%)	24 (75.0)	35 (68.6)	0.533
Education level, n (%)			
Primary school	8 (25.0)	12 (23.5)	
High school	10 (31.3)	14 (27.5)	
University	14 (43.8)	25 (49.0)	0.890
Marital status, n (%)			
Single	17 (53.1)	24 (47.1)	
Married	12 (37.5)	23 (45.1)	
Divorced/Widowed	3 (9.4)	4 (7.8)	0.790
BMI, kg/m²	30.42±5.43	29.32±6.74	0.451
Current smoker, n (%)	16 (51.6)	29 (56.9)	0.643
Medical comorbidity, n (%)	19 (59.4)	26 (51.0)	0.455
Clinical characteristics			
YMRS score	0.50±1.58	0.33±0.65	0.510
HAM-D score	2.54±2.27	2.78±2.30	0.646

BMI: Body mass index; YMRS: Young mania rating scale; HAM-D: Hamilton depression scale

Spearman's rank correlation coefficients were calculated to examine associations between serum lithium concentrations and PSQI, MEQ, and BRIAN scores. As a prespecified sensitivity analysis, lithium exposure was redefined according to therapeutic serum lithium concentrations. Patients with serum lithium levels ≥0.6 mmol/L were compared with all remaining participants (patients receiving non-lithium treatment and lithium-treated patients with serum concentrations <0.6 mmol/L). The adjusted regression analyses described above were repeated using this definition of therapeutic lithium exposure. In exploratory analyses, separate regression models were additionally performed for each BRIAN subdomain (sleep, activity, sociality, and eating patterns). A two-tailed p value<0.05 was considered statistically significant.

RESULTS

Participant Characteristics

A total of 83 euthymic patients with bipolar disorder were included, comprising 32 patients receiving lithium-containing treatment regimens and 51 receiving non-lithium treatment regimens. Patients receiving lithium were younger than those receiving non-lithium treatment (36.8±13.3 vs. 42.5±11.0 years, p=0.036). The two groups did not differ with respect to gender, educational attainment, marital status, body mass index, smoking status, medical comorbidity, or residual depressive and manic symptom severity (all p>0.05) (Table 1).

Table 2. Sleep and biological rhythm characteristics according to lithium treatment

	Lithium (n=32)	Non-lithium (n=50)	p	Effect size
PSQI total	7.00±5.45[1]	6.69±4.60[2]	0.788	0.06
MEQ total	51.72±7.39	51.45±9.54[2]	0.893	0.03
BRIAN total	35.09±8.12	38.90±13.42[2]	0.116	-0.33
BRIAN sleep	10.34±2.72	11.72±4.08	0.070	-0.38
BRIAN activity	9.75±3.36	10.74±4.58	0.263	-0.24
BRIAN social	7.47±2.21	8.04±3.45	0.364	-0.19
BRIAN eating	7.53±2.54	8.45±3.27[2]	0.183	-0.31

Values are presented as mean±standard deviation. [1]: n=31 because of missing data; [2]: n=49 because of missing data. PSQI: Pittsburgh sleep quality index; MEQ: Morningness–eveningness questionnaire; BRIAN: Biological rhythms interview of assessment in neuropsychiatry

Regarding current pharmacological treatment, combination pharmacotherapy was common in both the lithium and non-lithium groups [23 (71.9%) and 38 (74.5%), respectively; p=0.791]. Antipsychotic use was reported in 22 (68.8%) patients in the lithium group and 46 (90.2%) in the non-lithium group (p=0.013), while olanzapine/quetiapine use was comparable between groups [15 (46.9%) and 28 (54.9%), respectively; p=0.476]. Antidepressant use was reported in 5 (15.6%) patients in the lithium group and 15 (29.4%) in the non-lithium group (p=0.153). Benzodiazepine use was rare, with only two patients in the overall sample receiving clonazepam (1 mg/day), both of whom were in the non-lithium group. Anticonvulsant mood stabilizer use was more frequent in the non-lithium group [29 (56.9%)] than in the lithium group [2 (6.3%); p<0.001].

Sleep and Biological Rhythm Characteristics

In unadjusted analyses, no statistically significant between-group differences were observed in subjective sleep quality, chronotype, or biological rhythm measures (Table 2). However, patients receiving lithium consistently demonstrated numerically lower BRIAN total and subscale scores than those receiving non-lithium treatment.

PSQI total scores were significantly correlated with BRIAN total scores (Spearman's ρ=0.609, p<0.001) and with all BRIAN subdomains, including sleep (ρ=0.622), activity (ρ=0.556), social (ρ=0.498), and eating rhythms (ρ=0.408; all p<0.001).

Table 3. Multiple linear regression analyses examining the association between lithium treatment and biological rhythm measures[1]

	B (SE)	95% CI for B	β	p	Adjusted R^2
MEQ total	0.432 (1.990)	-3.54, 4.40	0.025	0.829	0.055
BRIAN total	-4.534 (2.251)	-9.02, -0.05	-0.193	0.048	0.341
BRIAN sleep	-1.653 (0.706)	-3.06, -0.25	-0.226	0.022	0.318
BRIAN activity	-1.001 (0.826)	-2.65, 0.65	-0.120	0.230	0.280
BRIAN social	-0.871 (0.654)	-2.18, 0.43	-0.140	0.187	0.184
BRIAN eating	-1.035 (0.659)	-2.35, 0.28	-0.169	0.121	0.167

[1]: Each model was adjusted for age, sex, and PSQI total score. Lithium treatment status was dummy-coded as 1 = lithium and 0 = non-lithium; therefore, negative coefficients indicate lower scores among patients receiving lithium. SE: Standard error; CI: Confidence interval; MEQ: Morningness–eveningness questionnaire; BRIAN: Biological rhythms interview of assessment in neuropsychiatry

Multivariable Regression Analyses

After adjustment for age, sex, and subjective sleep quality, lithium treatment was independently associated with lower biological rhythm disturbance, as reflected by lower BRIAN total scores (β=-0.193, p=0.048) (Table 3). No independent association was observed between lithium treatment and chronotype assessed using the MEQ (β=0.025, p=0.829). Exploratory analyses further demonstrated an independent association between lithium treatment and the BRIAN sleep domain (β=-0.226, p=0.022), whereas the remaining BRIAN subdomains were not significantly associated with lithium treatment.

Serum Lithium Concentrations

Among patients receiving lithium treatment, serum lithium concentrations were not significantly correlated with PSQI, MEQ, or BRIAN total scores (all p>0.05) (Table 4). Likewise, exploratory correlation analyses showed no significant associations between serum lithium concentrations and any of the individual BRIAN subdomains, including sleep, activity, social, or eating rhythms (all p>0.05).

Sensitivity Analyses

Of the 32 patients receiving lithium, 23 (71.9%) had serum lithium concentrations ≥0.6 mmol/L. To examine whether therapeutic lithium exposure was more strongly associated with biological rhythm disturbance than lithium treatment status alone, sensitivity analyses were performed after classifying lithium-treated patients according to a serum lithium concentration threshold of ≥0.6 mmol/L. Therapeutic lithium exposure was independently associated with lower BRIAN total scores (β=-0.244, p=0.011), whereas no association was observed with chronotype (p=0.554). Exploratory analyses additionally demonstrated significant associations with the BRIAN sleep, social, and eating domains, while the association with the activity domain did not reach statistical significance (p=0.079) (Table 5).

Table 4. Correlations between serum lithium levels and sleep and biological rhythm measures

	Spearman's ρ	p
PSQI total	0.036	0.850
MEQ total	-0.127	0.497
BRIAN total	-0.088	0.640
BRIAN sleep	0.046	0.807
BRIAN activity	-0.001	0.994
BRIAN social	-0.256	0.165
BRIAN eating	-0.101	0.590

PSQI: Pittsburgh sleep quality index; MEQ: Morningness eveningness questionnaire; BRIAN: Biological rhythms interview of assessment in neuropsychiatry

DISCUSSION

To our knowledge, this is one of the few clinical studies to compare sleep quality, chronotype, and biological rhythm disturbances between euthymic patients with bipolar disorder receiving lithium and those not receiving lithium while accounting for subjective sleep quality. The principal findings were fourfold. First, although unadjusted comparisons did not reveal significant between-group differences, lithium treatment was independently associated with lower biological rhythm disturbance after adjustment for age, sex, and subjective sleep quality. Second, lithium treatment was not associated with subjective sleep quality or chronotype. Third, serum lithium concentrations were not linearly associated with sleep or circadian measures. Finally, the association between lithium and biological rhythm disturbance became stronger when lithium exposure was defined according to therapeutic serum

Table 5. Associations between therapeutic lithium exposure and biological rhythm measures[1]

	B (SE)	95% CI for B	β	p	Adjusted R^2
MEQ total	-1.286 (2.163)	-5.60, 3.03	-0.067	0.554	0.057
BRIAN total	-6.229 (2.391)	-11.00, -1.46	-0.244	0.011	0.368
BRIAN sleep	-1.623 (0.773)	-3.16, -0.08	-0.204	0.039	0.308
BRIAN activity	-1.580 (0.886)	-3.35, 0.19	-0.174	0.079	0.301
BRIAN social	-1.541 (0.689)	-2.92, -0.17	-0.229	0.029	0.226
BRIAN eating	-1.504 (0.705)	-2.91, -0.10	-0.226	0.036	0.189

[1]: Adjusted for age, sex, and PSQI total score. Therapeutic lithium exposure was defined as a serum lithium concentration ≥0.6 mmol/L (n=23) and dummy-coded as 1 = serum lithium concentration ≥0.6 mmol/L and 0 = otherwise. SE: Standard error; CI: Confidence interval; MEQ: Morningness–eveningness questionnaire; BRIAN: Biological rhythms interview of assessment in neuropsychiatry

concentrations, suggesting that clinically effective lithium exposure may be more relevant than serum concentration itself.

Our findings suggest that the association between lithium treatment and biological rhythm disturbance may extend beyond differences in subjective sleep quality. Sleep and circadian rhythms are closely related but represent distinct physiological systems. Sleep quality primarily reflects the subjective experience of sleep and homeostatic sleep regulation, whereas biological rhythm measures capture the regularity and synchronization of daily behavioral rhythms, including sleep, activity, social interaction, and eating patterns. Although PSQI and BRIAN are conceptually overlapping measures, particularly because the BRIAN includes a sleep domain, the association between lithium treatment and BRIAN total score persisted after adjustment for PSQI. Notably, PSQI was significantly correlated with both BRIAN total and sleep-domain scores. Thus, the observed association with broader biological rhythm disturbance does not appear to be explained solely by differences in subjective sleep quality.

This interpretation is biologically plausible. Experimental studies have consistently demonstrated that lithium directly modulates the molecular circadian clock through inhibition of glycogen synthase kinase-3β (GSK-3β), thereby altering the phosphorylation and stability of several core clock proteins, including REV-ERBα, BMAL1, PER2, and CRY2. The molecular effects of lithium result in enhanced amplitude and stabilization of circadian rhythms in both animal models and cellular systems.[25] Unlike many psychotropic medications whose circadian effects are considered secondary to symptomatic improvement, lithium appears to exert direct effects on the molecular mechanisms regulating circadian rhythmicity. Our findings demonstrating an independent association between lithium treatment and biological rhythm disturbance in euthymic patients are consistent with these experimental observations. Nevertheless, the cross-sectional design does not allow us to determine whether the observed association reflects these molecular circadian mechanisms.

Clinical studies examining the relationship between lithium treatment and circadian functioning have yielded heterogeneous findings. Previous investigations using actigraphy, melatonin secretion, and self-report measures have generally suggested that lithium treatment is associated with greater stability of rest–activity rhythms, larger amplitude, and improved regularity of daily behavioral patterns, although findings have not been entirely consistent across methodologies and clinical populations.[26–28] These inconsistencies may partly reflect differences in mood state, illness stage, concomitant medications, and the circadian measures employed. Our findings extend this literature by demonstrating that, even among euthymic patients receiving routine maintenance treatment, lithium use was independently associated with lower biological rhythm disturbance after accounting for subjective sleep quality. The lack of association with chronotype further suggests that lithium may primarily improve circadian rhythm stability rather than alter intrinsic circadian preference.

Interestingly, lithium treatment was associated with lower BRIAN scores but not with chronotype as measured by the MEQ. Previous clinical studies investigating the association between lithium treatment and chronotype have yielded inconsistent results. Two relatively small cross-sectional studies found greater morningness among patients treated with lithium compared with patients taking other psychotropic medications, but a larger study of 525 patients found no difference in chronotype by lithium treatment status.[29–31] Consistent with the latter findings, we did not find an association between lithium treatment and MEQ scores, suggesting that lithium may affect the stability of biological rhythms rather than an individual's circadian preference. This difference is clinically relevant. Chronotype is a person's preferred timing of sleep and activity and is considered a relatively stable trait influ-

enced by genetic and environmental factors. In contrast, the BRIAN assesses the regularity of behavioral rhythms across multiple domains of daily life, including sleep, activity, social rhythms, and eating patterns.[23] Consequently, it may capture changes in the organization of daily rhythms that are not necessarily reflected by measures of chronotype. The findings suggest that lithium treatment may be more closely associated with the stability and regularity of biological rhythms than with circadian preference. This pattern is consistent with the hypothesis that lithium treatment may be related to circadian rhythm stability rather than chronotype modification.

Another notable finding was that serum lithium concentrations were not linearly associated with biological rhythm measures. However, when patients were classified according to a therapeutic serum lithium threshold (≥0.6 mmol/L), the association with biological rhythm disturbance became stronger and extended to several BRIAN domains. Although exploratory, this observation raises the possibility that associations with biological rhythm stability may be more apparent at therapeutically relevant lithium concentrations than across serum lithium concentrations considered as a continuous measure. Such a threshold effect would also be consistent with clinical practice, in which therapeutic efficacy is generally expected once serum lithium concentrations reach the recommended maintenance range.

Interestingly, the sensitivity analyses suggested that the association between lithium treatment and biological rhythms was not confined to a single behavioral domain but extended across sleep, social, and eating rhythms, with a similar trend observed for activity rhythms. Although these analyses are exploratory, this pattern may suggest that lithium treatment is associated with the overall temporal organization of daily behaviors rather than isolated effects on sleep alone. This interpretation is in line with the social zeitgeber hypothesis, which posits that the regularity of daily routines, including sleep, meals, physical activity, and social interactions, plays a central role in maintaining circadian stability and reducing vulnerability to mood episodes in bipolar disorder. Rather than being associated with a particular circadian output, lithium treatment may be associated with greater synchronization across multiple behavioral rhythms and greater stability of the circadian system as a whole. Nevertheless, this interpretation should be considered preliminary and requires investigation in prospective longitudinal studies.

Our study has several strengths. Participants were restricted to euthymic patients, minimizing the confounding effect of acute mood symptoms on sleep and circadian measures. In addition, we distinguished subjective sleep quality from broader biological rhythm disturbance and evaluated serum lithium concentrations in addition to treatment status. Furthermore, by adjusting for subjective sleep quality, we specifically examined whether the association between lithium and biological rhythms was independent of its potential effects on sleep, providing a more direct evaluation of circadian rhythm.

Several limitations should also be acknowledged. First, the cross-sectional design precludes causal inference, and it cannot be determined whether lithium treatment improves biological rhythms or whether patients with more stable biological rhythms are more likely to remain on lithium treatment. Second, many patients in the lithium group were also receiving additional psychotropic medications as part of routine clinical care. Although this reflects real-world clinical practice and enhances the generalizability of our findings, the independent contribution of lithium cannot be completely separated from that of concomitant medications. Importantly, several psychotropic agents, including atypical antipsychotics, anticonvulsants, and antidepressants, have also been reported to influence sleep architecture and circadian rhythm regulation. Therefore, residual confounding related to concomitant pharmacotherapy cannot be excluded. Third, although adjustment for PSQI allowed us to examine whether the association with biological rhythm disturbance persisted after accounting for subjective sleep quality, some conceptual overlap between PSQI and BRIAN should be acknowledged because the BRIAN includes a sleep domain. Moreover, biological rhythms and sleep were assessed using validated subjective and interview-based measures rather than objective circadian measures, such as actigraphy, dim-light melatonin onset, or core body temperature. Finally, the relatively small sample size may have limited the statistical power to detect small between-group differences, especially in the unadjusted analyses.

CONCLUSION

In conclusion, lithium treatment was independently associated with lower biological rhythm disturbance, but not with subjective sleep quality or chronotype, in euthymic patients with bipolar disorder. Associations with biological rhythm disturbance were more pronounced when lithium exposure was defined using a therapeutic serum concentration threshold. Together, these findings are consistent with the hypothesis that lithium treatment may be associated with greater biological rhythm stability beyond differences in subjective sleep quality. Longitudinal studies incorporating objective circadian measures are needed to determine the temporal and potentially mechanistic nature of this relationship.

Disclosures

Ethics Committee Approval: The study was approved by the Istanbul Medeniyet University Göztepe Prof. Dr. Süleyman Yalçın City Hospital Clinical Research Ethics Committee (No: 2025/0208, Date: 09/10/2025).

Informed Consent: All participants provided written informed consent before enrolment.

Conflict of Interest Statement: The authors declare that there is no conflict of interest.

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

Use of AI for Writing Assistance: The authors declare that they have no competing interests. The authors confirm that this manuscript is original and does not contain plagiarized material. The authors used generative artificial intelligence (ChatGPT, OpenAI) for language refinement and editing purposes only.

Author Contributions: Concept – A.K., F.B.P., Z.Ç.; Design – A.K., F.B.P.; Supervision – Z.Ç.; Funding – A.K.; Materials – P.E., H.Ş.; Data collection and/or processing – P.E., H.Ş.; Data analysis and/or interpretation – A.K.; Literature search – A.K., F.B.P., Z.Ç.; Writing – A.K.; Critical review – A.K., F.B.P., Z.Ç., P.E., H.Ş.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Oliva V, Fico G, De Prisco M, Gonda X, Rosa AR, Vieta E. Bipolar disorders: an update on critical aspects. Lancet Reg Health Eur 2024;48:101135. [CrossRef]
2. Swidzinski S, Tsapekos D, Swidzinska P, Zhang W, Zheng M, Millgate E, et al. Domain-specific cognitive function in euthymic bipolar disorder: a systematic review and meta-analysis. Psychol Med 2025;55:e336. [CrossRef]
3. Grover S, Avasthi A, Chakravarty R, Dan A, Chakraborty K, Neogi R, et al. Residual symptoms in bipolar disorders: Findings from the bipolar Disorder course and outcome study from India (BiD-CoIN study). Psychiatry Res 2021;302:113995. [CrossRef]
4. Gitlin MJ, Miklowitz DJ. The difficult lives of individuals with bipolar disorder: A review of functional outcomes and their implications for treatment. J Affect Disord 2017;209:147–54. [CrossRef]
5. Ng TH, Chung KF, Ho FY, Yeung WF, Yung KP, Lam TH. Sleep-wake disturbance in interepisode bipolar disorder and high-risk individuals: a systematic review and meta-analysis. Sleep Med Rev 2015;20:46–58. [CrossRef]
6. De la Fuente-Tomás L, Sierra P, Sanchez-Autet M, García-Blanco A, Safont G, Arranz B, et al. Sleep disturbances, functioning, and quality of life in euthymic patients with bipolar disorder. Psychiatry Res 2018;269:501–7. [CrossRef]
7. Deboer T. Sleep homeostasis and the circadian clock: Do the circadian pacemaker and the sleep homeostat influence each other's functioning? Neurobiol Sleep Circadian Rhythms 2018;5:68–77. [CrossRef]
8. Bass J, Lazar MA. Circadian time signatures of fitness and disease. Science 2016;354:994–9. [CrossRef]
9. Walker WH, 2nd, Walton JC, DeVries AC, Nelson RJ. Circadian rhythm disruption and mental health. Transl Psychiatry 2020;10:28. [CrossRef]
10. Merikangas KR, Swendsen J, Hickie IB, Cui L, Shou H, Merikangas AK, et al. Real-time mobile monitoring of the dynamic associations among motor activity, energy, mood, and sleep in adults with bipolar disorder. JAMA Psychiatry 2019;76:190–8. [CrossRef]
11. Wood J, Birmaher B, Axelson D, Ehmann M, Kalas C, Monk K, et al. Replicable differences in preferred circadian phase between bipolar disorder patients and control individuals. Psychiatry Res 2009;166:201–9. [CrossRef]
12. Romo-Nava F, Blom TJ, Cuellar-Barboza AB, Winham SJ, Colby CL, Nunez NA, et al. Evening chronotype as a discrete clinical subphenotype in bipolar disorder. J Affect Disord 2020;266:556–62. [CrossRef]
13. Cruickshank HK, Varela RB, Truong TTT, Walder K, Berk M, Tye SJ. Melatonin in bipolar disorder. J Affect Disord 2026;398:120871. [CrossRef]
14. Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. Nat Genet 2021;53:817–29.
15. Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Bond DJ, Frey BN, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) 2018 guidelines for the management of patients with bipolar disorder. Bipolar Disord 2018;20:97–170.
16. Rohr KE, McCarthy MJ. The impact of lithium on circadian rhythms and implications for bipolar disorder pharmacotherapy. Neurosci Lett 2022;786:136772. [CrossRef]
17. Karadağ F, Oral ET, Aran Yalçın F, Erten E. Reliability and validity of Turkish translation of young mania rating scale. Turk Psikiyatri Derg 2001;13:107–14. [Turkish]
18. Akdemir A, Türkçapar MH, Orsel SD, Demirergi N, Dag I, Ozbay MH. Reliability and validity of the Turkish version of the Hamilton Depression Rating Scale. Compr Psychiatry 2001;42:161–5. [CrossRef]
19. Buysse DJ, Reynolds CF, 3rd, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213. [CrossRef]
20. Ağargün MY. Validity and reliability of Turkish version of Pittsburgh sleep quality index. Türk Psikiyatri Dergisi 1996;7:107.
21. Horne JA, Ostberg O. A self-assessment questionnaire to determine morningness-eveningness in human circadian rhythms. Int J Chronobiol 1976;4:97–110. [CrossRef]
22. Agargun MY, Cilli AS, Boysan M, Selvi Y, Gulec M, Kara H. Turkish version of morningness-eveningness questionnaire (MEQ). Sleep Hypn 2007;9:16.
23. Giglio LM, Magalhães PV, Andreazza AC, Walz JC, Jakobson L, Rucci P, et al. Development and use of a biological rhythm interview. J Affect Disord 2009;118:161–5. [CrossRef]
24. Aydemir O, Akkaya C, Altinbas K, Kora K, Suculluoglu Dikici D, Akdeniz F, et al. Reliability and validity of Turkish version of biological rhythms interview of assessment in neuropsychiatry. Anatol J Psychiatry 2012;13:256–61.
25. Li J, Lu W-Q, Beesley S, Loudon AS, Meng Q-J. Lithium impacts on the amplitude and period of the molecular circadian clockwork. PLoS One 2012;7:e33292. [CrossRef]
26. Scott J, Hennion V, Meyrel M, Bellivier F, Etain B. An ecological study of objective rest-activity markers of lithium response in bipolar-I-disorder. Psychol Med 2022;52:2281–9. [CrossRef]
27. Hwang JY, Choi J-W, Kang S-G, Hwang SH, Kim SJ, Lee YJ. Comparison of the effects of quetiapine XR and lithium monotherapy on actigraphy-measured circadian parameters in patients with bipolar II depression. J Clin Psychopharmacol 2017;37:351–4. [CrossRef]

28. Xu N, Shinohara K, Saunders KEA, Geddes JR, Cipriani A. Effect of lithium on circadian rhythm in bipolar disorder: A systematic review and meta-analysis. Bipolar Disord 2021;23:445–53. [CrossRef]
29. Geoffroy PA, Samalin L, Llorca PM, Curis E, Bellivier F. Influence of lithium on sleep and chronotypes in remitted patients with bipolar disorder. J Affect Disord 2016;204:32–9. [CrossRef]
30. Dopierała E, Chrobak A, Tereszko A, Rybakowski J. Lithium influence on circadian rhythm assessed by the Composite Scale of Morningness in remitted bipolar patients. Pharmacother Psychiatry Neurol 2017;22:9–20.
31. Kanagarajan K, Gou K, Antinora C, Buyukkurt A, Crescenzi O, Beaulieu S, et al. Morningness-Eveningness questionnaire in bipolar disorder. Psychiatry Res 2018;262:102–7. [CrossRef]

DOI: 10.14744/cm.2026.52297
Compreh Med 2026;18(Suppl. 1):103-111

RESEARCH ARTICLE

Clinical, Microbiological and Treatment Characteristics of Spondylodiscitis: A Six-Year Single-Centre Neurosurgery Experience from a Brucellosis-Endemic Region of Türkiye

Salim Katar[1], Adil Can Karaoğlu[1], Muhammed Fatih Sarı[1], Mehmet Ali Tüz[2], Rukiye İnan Sarıkaya[2], Emre Fidan[3], Hüseyin Utku Adilay[1]

[1]Department of Neurosurgery, Balıkesir University Faculty of Medicine, Balıkesir, Türkiye
[2]Department of Infectious Diseases and Clinical Microbiology, Balıkesir University Faculty of Medicine, Balıkesir, Türkiye
[3]Balıkesir University Faculty of Medicine, Balıkesir, Türkiye

ABSTRACT

Objective: The microbiological spectrum of spondylodiscitis (SD) varies markedly by region; in *Brucella*-endemic settings, the etiological profile differs from that of Western cohorts, yet contemporary data from neurosurgical practice in Türkiye are limited. We aimed to characterize SD managed in a tertiary neurosurgical service in *Brucella*-endemic Western Türkiye and to identify variables associated with prolonged intravenous antibiotic therapy.

Materials and Methods: Consecutive adult patients with SD treated between January 2020 and December 2025 at a tertiary academic medical center were retrospectively analyzed. Group comparisons were performed using the Mann–Whitney U test, chi-square test, and Fisher's exact test, as appropriate; correlations were assessed using Spearman's rho. A two-sided $p<0.05$ was considered statistically significant.

Results: Fifty-six patients were included (median age, 66 years [IQR, 59–71]; 53.6% male; lumbar involvement, 55.4%). Surgery was performed in 49 patients (87.5%; 95% CI, 76.4–93.8%), and an abscess was identified in 35 (62.5%; 95% CI, 49.4–74.0%). A causative organism was isolated in 29 patients (51.8%); *Brucella* spp. (14.3%) and *Mycobacterium tuberculosis* (7.1%) predominated over *Staphylococcus aureus* (3.6%). The median duration of intravenous antibiotic therapy was 36 days (IQR, 13–52). Surgical management ($p=0.002$) and abscess presence ($p=0.008$) were associated with longer antibiotic therapy; abscess presence was strongly associated with surgery (Fisher's exact $p=0.009$). Pyogenic SD required approximately twice the duration of antibiotic therapy compared with granulomatous SD ($p=0.021$).

Conclusion: In this *Brucella*-endemic neurosurgical cohort, granulomatous pathogens predominated over *S. aureus*, and abscess presence was the central determinant of both operative intervention and prolonged antimicrobial therapy.

Keywords: Antibiotic treatment, brucellosis, neurosurgery, spinal abscess, spondylodiscitis, vertebral osteomyelitis

How to cite this article: Katar S, Karaoğlu AC, Sarı MF, Tüz MA, İnan Sarıkaya R, Fidan E, Adilay HU. Clinical, Microbiological and Treatment Characteristics of Spondylodiscitis: A Six-Year Single-Centre Neurosurgery Experience from a Brucellosis-Endemic Region of Türkiye. Compreh Med 2026;18(Suppl. 1):103-111.

INTRODUCTION

Spondylodiscitis (SD) refers to an infection of the intervertebral disc and adjacent vertebral endplates and frequently spreads into the paravertebral or epidural space.[1,2] The reported incidence has increased significantly over the last twenty years, reaching 14.4 cases per 100,000 people in some European registries; factors contributing to this increase include an ageing population, the increasing prevalence of spinal instrumentation, and a rising burden of comorbidities.[2,3] Despite advances in imaging and microbiological diagnostic methods, SD remains associated with significant morbidity; in-hospital mortality remains within the range of 2–17%,

Address for Correspondence: Salim Katar, Department of Neurosurgery, Balıkesir University Faculty of Medicine, Balıkesir, Türkiye
E-mail: salimkatar@gmail.com **ORCID ID:** 0000-0002-8311-1122

Received date: 16.06.2026
Revised date: 05.09.2026
Accepted date: 17.09.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

and permanent disability is common.[3,4] The reasons for this rising incidence are multifactorial and include the ageing of European and Mediterranean populations with an associated burden of comorbidities (diabetes mellitus, immunosuppression, chronic haemodialysis), the broader use of spinal instrumentation and percutaneous interventional procedures, the spread of intravenous drug use, and improved diagnostic recognition with the wider availability of magnetic resonance imaging. Healthcare-associated and postsurgical spondylodiscitis now represents a growing proportion of all cases in centres performing high volumes of spinal surgery, with reported rates of 1–4% following instrumented procedures.

The microbiological spectrum of SD exhibits marked geographical variations. In Western series, pyogenic pathogens, predominantly *Staphylococcus aureus*, account for 20–80% of cases.[1,5] In contrast, in endemic regions of the Mediterranean and the Middle East, granulomatous pathogens, particularly *Brucella* spp. and *Mycobacterium tuberculosis*, may constitute a large proportion of the identified organisms.[6–8] Türkiye occupies an intermediate epidemiological position: brucellosis remains endemic in provinces with livestock farming, and spinal involvement is observed in up to 50% of focal brucellosis cases.[6,9] In parallel, multicentre series in Türkiye have documented that *Brucella* is the most frequently identified pathogen in spondylodiscitis cohorts.[7,10]

Treatment requires a multidisciplinary approach combining long-term antimicrobial therapy with surgery in cases of neurological deficit, structural instability, abscess formation, or failure of conservative treatment.[1,5] The 2015 Infectious Diseases Society of America (IDSA) guidelines recommend a minimum of six weeks of antimicrobial therapy for native vertebral osteomyelitis, supported by the landmark randomised trial by Bernard et al.,[11] which demonstrated the non-inferiority of six-week treatment compared with twelve-week treatment. However, in real-world practice, the duration of antibiotic therapy often significantly exceeds these reference values, particularly in cases involving complicating factors such as abscesses or instrumentation.[12,13] The clinical variables leading to such treatment intensification have not been sufficiently characterised in neurosurgical practice in Türkiye.

In this study, we aimed to characterise the clinical, microbiological, and treatment characteristics of SD treated over a six-year period in a tertiary neurosurgery department in a Western Anatolian province with a high concentration of livestock farming and to identify variables associated with prolonged intravenous antibiotic therapy and surgical treatment.

MATERIALS and METHODS

Study Design and Setting

We conducted a single-centre retrospective observational study at the Departments of Neurosurgery and Infectious Diseases of a tertiary academic medical centre in Western Anatolia. Consecutive adult patients (aged ≥18 years) admitted with a diagnosis of spondylodiscitis between 1 January 2020 and 31 December 2025 were identified from the hospital's electronic medical records using ICD-10 codes M46.2–M46.5 and M49. The reporting of this study adheres to the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement for cohort studies.

The diagnosis of SD was based on the following: (i) a consistent clinical presentation (back pain±fever±neurological deficit); (ii) characteristic findings on contrast-enhanced magnetic resonance imaging (endplate destruction, narrowing of the disc space or signal changes, paravertebral or epidural soft tissue involvement); and (iii) confirmation by microbiological isolation from blood, intraoperative tissue, or image-guided biopsy, where possible.[1,5]

Variables and Definitions

The following variables were summarised: age, sex, presence of fever on admission, level of anatomical involvement, aetiology (primary haematogenous vs. secondary [postoperative or post-procedural]), causative organism, presence of a paravertebral or epidural abscess on the first MRI, surgical treatment, and total duration of intravenous antibiotic therapy (in days).

Aetiology was classified as primary if it arose without prior spinal intervention and as secondary if it developed following spinal surgery, a percutaneous procedure, or an adjacent infection. For analytical purposes, the microbiological spectrum was further categorised as culture-negative, pyogenic (Gram-positive cocci, Gram-negative bacilli, enterococci), granulomatous (*Brucella* spp., *M. tuberculosis*), or fungal. Spinal levels were categorised as lumbar, lumbosacral, thoracic, thoracolumbar, or other. Operative management was defined as any surgical intervention performed for spondylodiscitis during hospitalisation, including open or minimally invasive decompression, debridement, instrumented or non-instrumented fusion, and combined approaches; the choice of procedure was at the discretion of the treating neurosurgeon based on the level of involvement, structural integrity, neurological status, and abscess burden.

Statistical Analysis

The distribution of continuous variables was assessed using the Shapiro–Wilk test; none were normally distributed and were therefore reported as the median, interquartile range (IQR), and full range. Categorical variables were presented as numbers (%). For group comparisons, the Mann–Whitney U test was used for continuous data and the Pearson chi-square or Fisher's exact test for categorical data, as appropriate. Spearman's rho was used for correlation analysis. A two-sided $p<0.05$ was considered statistically significant. Analyses were performed using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA).

Ethics

The study was conducted in accordance with the Declaration of Helsinki. Ethical approval was obtained from the Institutional Clinical Research Ethics Committee (Approval No.: 2026/9-7; Date: 05/05/2026). Due to the retrospective design of the study, the requirement for written informed consent was waived.

RESULTS

Demographic and Clinical Characteristics

Fifty-six patients met the inclusion criteria. The median age was 66 years (IQR 59–71; range 27–84); 30 (53.6%) were male. Most patients (37/56, 66.1%) were aged 61–75 years; this reflects the highest incidence, which is well established among older adults. Consistent with the typically indolent clinical presentation of SD, fever was detected in only 8 patients (14.3%; 95% CI 7.4–25.7%) on admission. Demographic and clinical characteristics are summarised in Table 1.

The most commonly affected region was the lumbar spine (31 patients, 55.4%), followed by the thoracic spine (11, 19.6%), the lumbosacral junction (9, 16.1%), and the thoracolumbar junction (4, 7.1%). One patient had multilevel involvement that could not be classified into these categories. Secondary (post-procedural) aetiology was predominant (43, 76.8%); primary haematogenous SD was detected in 10 patients (17.9%), and the route of transmission could not be reliably determined in 3 patients (5.4%).

Microbiological Profile

The causative microorganism was identified in 29 of the 56 patients (51.8%; 95% CI 39.0–64.3%); in the remaining 27 patients (48.2%; 95% CI 35.7–61.0%), blood cultures and, where performed, intraoperative tissue samples were obtained, but culture results were negative. *Brucella* spp. was the most frequently isolated pathogen (8 patients, 14.3%; 95% CI 7.4–25.7%), followed by *M. tuberculosis* (4, 7.1%). Pyogenic bacteria accounted for a total of 16 cases (28.6%); these included coagulase-negative staphylococci (3, 5.4%), MRSA/methicillin-resistant coagulase-negative staphylococci (3, 5.4%), *S. aureus* (2, 3.6%), *S. epidermidis* (2, 3.6%), and single cases of *Klebsiella pneumoniae, Enterobacter spp., Proteus* spp., *Acinetobacter* spp., *Salmonella* spp., and *Enterococcus* spp. A single fungal case (*Candida* spp., 1.8%) was identified. The full distribution is shown in Table 2 and depicted graphically in Figure 1. The microbiological yield of 51.8% in this cohort is comparable to the ranges reported from European tertiary centres (50–70%) and recent single-centre series from Türkiye, although direct comparison is complicated by differences in sampling strategy, biopsy timing, and the extent of empirical antibiotic exposure before tissue acquisition.

Table 1. Demographic and clinical characteristics of the study population (n=56)

Variable	Value
Age (years), median (IQR), [range]	66 (59–71),(27–84)
Age category, n (%)	
≤40	4 (7.1)
41–60	11 (19.6)
61–75	37 (66.1)
>75	4 (7.1)
Male sex, n (%)	30 (53.6)
Fever at admission, n (%)	8 (14.3)
Aetiology, n (%)	
Primary	10 (17.9)
Secondary	43 (76.8)
Uncertain	3 (5.4)
Anatomical level, n (%)	
Lumbar	31 (55.4)
Lumbosacral	9 (16.1)
Thoracic	11 (19.6)
Thoracolumbar	4 (7.1)
Other	1 (1.8)
Abscess on MRI, n (%)	35 (62.5)
Surgical treatment, n (%)	49 (87.5)
Duration of IV antibiotics (days)[a], median (IQR), [range]	36 (13–52),(6–94)

[a]: Duration available for 55 patients. IQR: Interquartile range; MRI: Magnetic resonance imaging; IV: Intravenous

Table 2. Microbiological distribution of identified pathogens (n=56)

Category / Organism	n	%
Granulomatous	12	21.4
Brucella spp.	8	14.3
Mycobacterium tuberculosis	4	7.1
Pyogenic Gram-positive	10	17.9
Coagulase-negative staphylococci	3	5.4
MRSA / MRCoNS	3	5.4
Staphylococcus aureus	2	3.6
S. epidermidis	2	3.6
Pyogenic Gram-negative	5	8.9
Klebsiella pneumoniae	1	1.8
Enterobacter spp.	1	1.8
Proteus spp.	1	1.8
Acinetobacter spp.	1	1.8
Salmonella spp.	1	1.8
Enterococci	1	1.8
Fungi (*Candida* spp.)	1	1.8
Culture negative	27	48.2

MRSA: Methicillin-resistant *S. aureus*; MRCoNS: Methicillin-resistant coagulase-negative staphylococci

Treatment Modalities and Duration of Antibiotic Therapy

Surgical treatment comprising decompression, debridement, and/or instrumented fusion was performed in 49 patients (87.5%; 95% CI 76.4–93.8%) based on clinical indications; the remaining 7 patients (12.5%) were treated conservatively with long-term antibiotic therapy and brace use. The median duration of intravenous antibiotic therapy was 36 days (IQR 13–52; range 6–94) among the 55 patients for whom treatment duration was documented; data for this variable were missing for one patient.

Factors Associated with the Duration of Antibiotic Therapy

Two factors were significantly associated with a longer duration of intravenous antibiotic treatment (Table 3). Patients who underwent surgery received significantly longer treatment than those who received conservative treatment (median 38 days [IQR 17–54] vs. 8 days [IQR 7–10]; U=262, p=0.002). Similarly, patients in whom an abscess was detected on imaging received longer treatment than those without an abscess (39 days [IQR 22–58] vs. 16 days [IQR 11–37]; U=502, p=0.008) (Fig. 2a,b). In contrast, aetiology (primary vs. secondary; p=0.347), fever on admission (p=0.830), and age (Spearman's rho=0.105, p=0.446) were not associated with the duration of antibiotic treatment.

Table 3. Factors associated with the duration of intravenous antibiotic therapy

Variable	n	Median (IQR) days	Test statistic	p
Surgical treatment			U=262	**0.002**
Yes	49	38 (17–54)		
No	6	8 (7–10)		
Abscess on MRI			U=502	**0.008**
Present	35	39 (22–58)		
Absent	20	16 (11–37)		
Aetiology			U=169	0.347
Primary	10	20 (12–42)		
Secondary	42	38 (14–54)		
Fever on admission			U=198	0.830
Present	8	34 (10–66)		
Absent	47	36 (14–46)		
Age (Spearman)	55	rho=0.105		0.446

Bolded values indicate statistical significance (p<0.05). All comparisons, except for age (Spearman correlation), were performed using the Mann–Whitney U test. IQR: Interquartile range; MRI: Magnetic resonance imaging

Abscess as a Determinant of Surgical Treatment

An abscess was identified on imaging in 35 of the 56 patients. Of these, 34 (97.1%) underwent surgery, whereas only 15 of the 21 patients without an abscess (71.4%) required surgical intervention (Table 4; Fig. 3). The association between the presence of an abscess and surgical treatment was statistically significant (Fisher's exact test: odds ratio=13.60, p=0.009).

Granulomatous and Pyogenic Spondylodiscitis

A comparative analysis of culture-positive cases revealed clinically significant differences (Table 5). Patients with pyogenic SD required significantly longer courses of intravenous antibiotic treatment than those with granulomatous SD (median 43 days [IQR 36–60] vs. 21 days [IQR 15–39]; p=0.021; Fig. 2c). All pyogenic cases (16/16, 100%) underwent surgery, whereas 9 of 12 granulomatous cases (75.0%) underwent surgery. Thoracic involvement was disproportionately common in the granulomatous group (5/12, 41.7%) compared with the pyogenic group (2/16, 12.5%); this is consistent with the classic tendency of *M. tuberculosis* to affect the thoracic spine.

DISCUSSION

In this six-year single-centre series involving 56 patients with spondylodiscitis treated at a tertiary neurosurgery de-

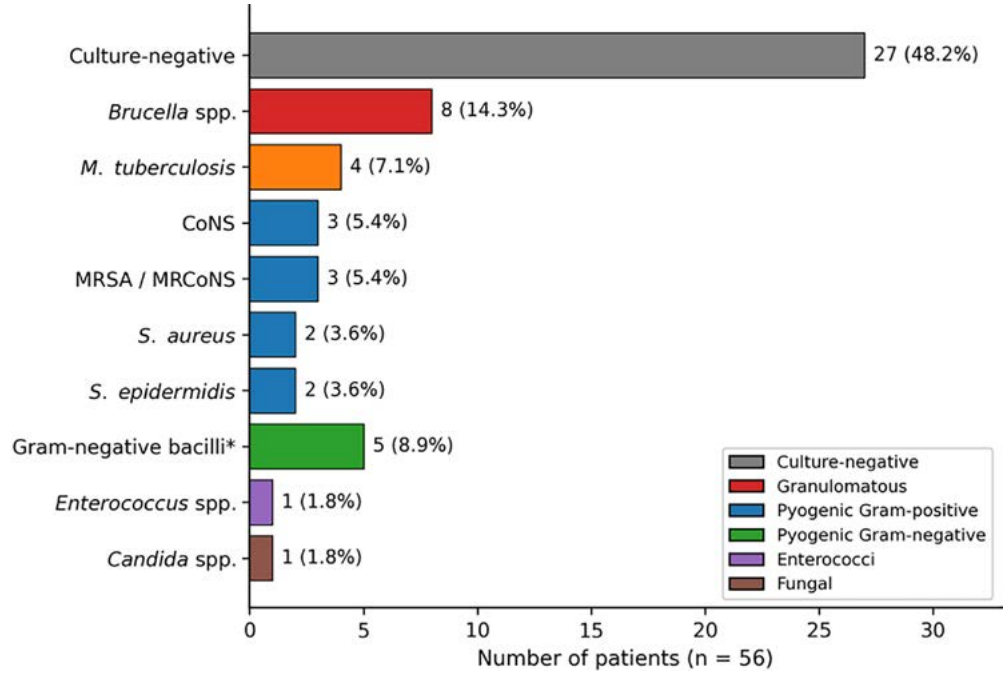


Figure 1. Microbiological distribution of pathogens identified in the study cohort (n=56). The bars are colour-coded according to pathogen category: culture-negative (grey), granulomatous (red/orange), pyogenic Gram-positive (blue), pyogenic Gram-negative (green), enterococci (purple) and fungal (brown). *Brucella* spp. was the most frequently identified single pathogen. *Gram-negative bacilli include *Klebsiella pneumoniae, Enterobacter* spp., *Proteus* spp., *Acinetobacter* spp. and *Salmonella* spp. (one case each)

CoNS: Coagulase-negative staphylococci; MRSA: Methicillin-resistant S. aureus; MRCoNS: Methicillin-resistant coagulase-negative staphylococci

partment in western Türkiye, three key findings emerged. First, the microbiological profile was dominated by granulomatous pathogens (*Brucella* spp. and *M. tuberculosis*); these accounted for 21.4% of all cases and were nearly five times more common than *S. aureus*. Second, the presence of an abscess was the central clinical determinant of treatment intensity and was strongly associated with both surgical intervention and prolonged intravenous antibiotic therapy. Third, despite modern diagnostic methods, the culture-negative rate remained substantial (48.2%), highlighting an ongoing diagnostic challenge in clinical practice.

Our microbiological distribution differs markedly from Western European cohorts, where *S. aureus* is typically predominant and granulomatous aetiologies are rare.[1–5,12] In a recent 14-year Italian series by Conti et al.,[14] *S. aureus* accounted for nearly half of the identified pathogens, and complications such as endocarditis and paravertebral abscesses were frequently observed. In contrast, in a comparative study from Türkiye conducted by Bal et al.[7] in Thrace involving 109 patients, brucellar spondylodiscitis accounted for a significant proportion of all cases and exhibited distinct epidemiological and clinical characteristics compared with pyogenic disease. The 14.3% *Brucella* rate in our study is consistent with reports from other Mediterranean basin countries[6,9] and confirms that, despite ongoing public health efforts, brucellosis remains a leading endemic cause of SD in Türkiye. It is worth noting that although brucellosis shows particularly high transmission in livestock-farming provinces such as ours,

Table 4. Cross-tabulation of abscess presence and surgical treatment

Abscess status	Surgery: Yes	Surgery: No	Total
Abscess present	34 (97.1%)	1 (2.9%)	35
Abscess absent	15 (71.4%)	6 (28.6%)	21
Total	49	7	56

Fisher's exact test (two-sided): odds ratio=13.60, p=0.009. Pearson chi-square was inappropriate due to expected cell count <5

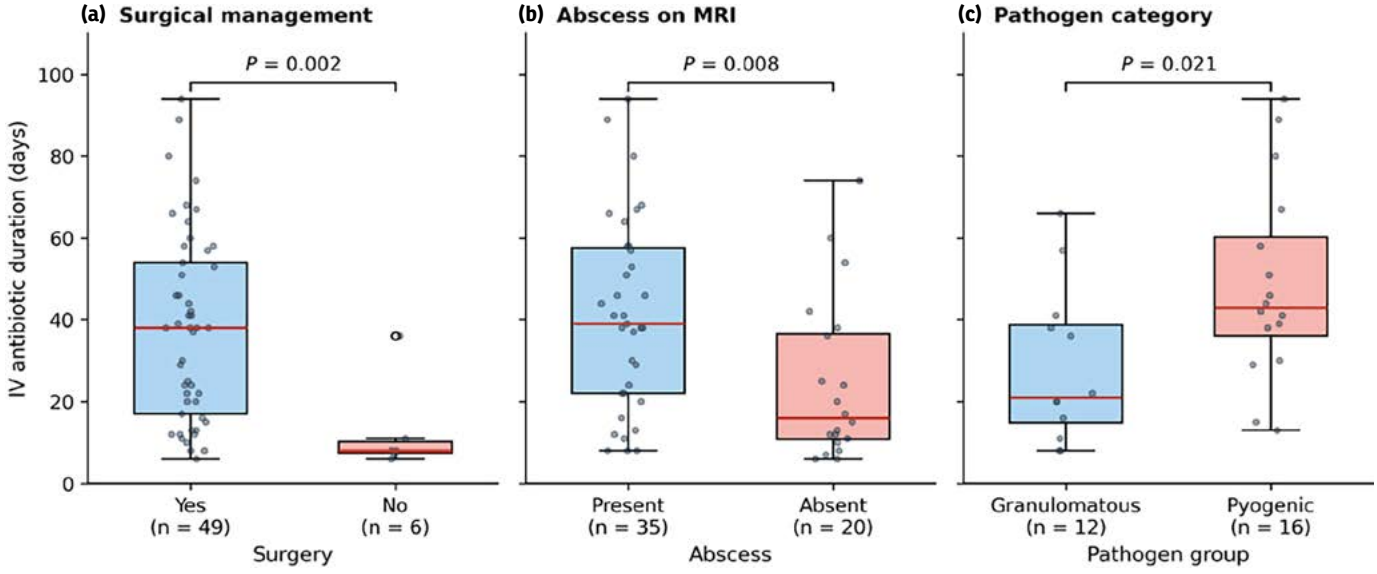


Figure 2. Duration of intravenous antibiotic therapy classified by **(a)** surgical treatment, **(b)** presence of an abscess on magnetic resonance imaging, and **(c)** pathogen category. Boxes show the interquartile range (red line); bars extend up to 1.5 × IQR. Individual patient values are shown as semi-transparent dots. All three comparisons were found to be statistically significant using the Mann–Whitney U test

IQR: İnterquartile range

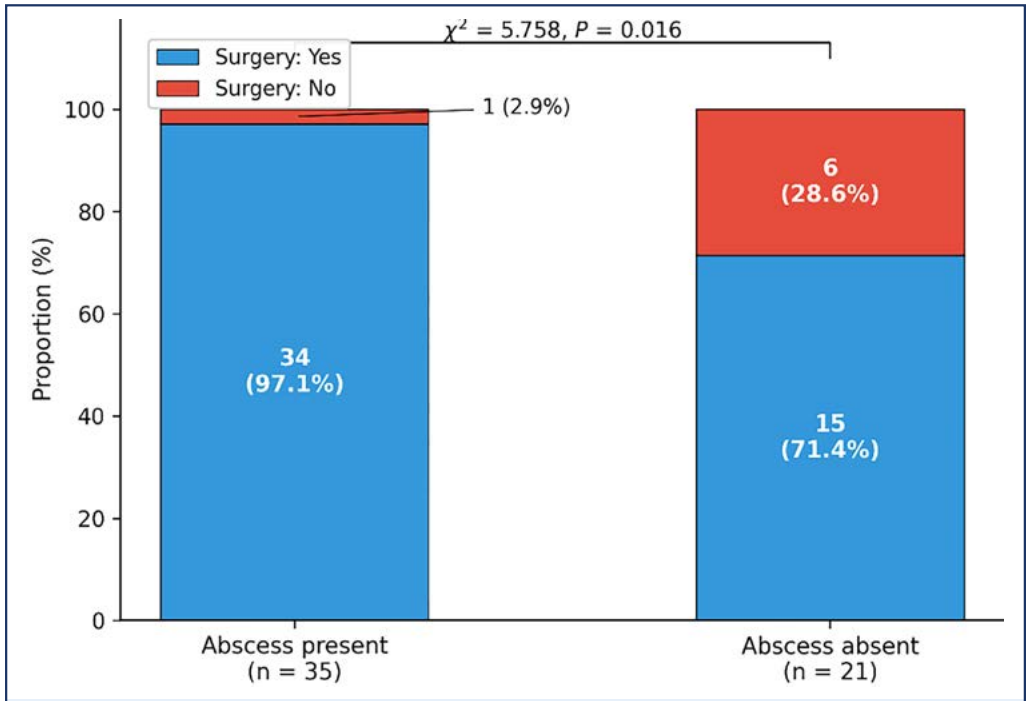


Figure 3. Distribution of surgical and conservative treatment by abscess status. Of the 35 patients in whom an abscess was detected on magnetic resonance imaging, 34 (97.1%) underwent surgical treatment, whereas of the 21 patients in whom no abscess was detected, 15 (71.4%) underwent surgical treatment. This association is statistically significant (Fisher's exact p=0.009)

Table 5. Comparison of granulomatous and pyogenic spondylodiscitis

Variable	Granulomatous (n=12)	Pyogenic (n=16)	p
Age (years), median (IQR)	61 (55–67)	64 (60–71)	0.295
Male sex, n (%)	7 (58.3)	8 (50.0)	0.718
Thoracic involvement, n (%)	5 (41.7)	2 (12.5)	0.103
Lumbar involvement, n (%)	4 (33.3)	10 (62.5)	0.252
Abscess on MRI, n (%)	9 (75.0)	14 (87.5)	0.624
Surgical treatment, n (%)	9 (75.0)	16 (100.0)	0.067
Duration of IV antibiotics (days), median (IQR)	21 (15–39)	43 (36–60)	**0.021**

Granulomatous = *Brucella* spp. + *M. tuberculosis*. Pyogenic = all bacterial pathogens other than granulomatous. Continuous variables were analysed using the Mann–Whitney U test; categorical variables were analysed using Fisher's exact test (two-sided). IQR: Interquartile range; MRI: Magnetic resonance imaging; IV: Intravenous

the disease is not confined to these regions and continues to be reported, albeit at a lower incidence, across Türkiye as a whole. Importantly, granulomatous pathogens may require entirely different antimicrobial strategies, typically rifampicin- or streptomycin-containing regimens combined with doxycycline for brucellosis[9,15] and standard four-drug antituberculosis therapy for spinal tuberculosis,[8] underscoring the importance of early aetiological clarification.

The high culture-negative rate of 48.2% in our cohort is consistent with the wide range reported in the literature (30–70%).[5,16] Several factors likely contribute to this finding, including prior exposure to empirical antibiotics before culture sampling, the slow progression of granulomatous infections caused by difficult-to-culture organisms, and the diagnostic yield gradient whereby blood cultures (≈40–60%) and percutaneous biopsies (≈60–70%) consistently underperform compared with intraoperative tissue cultures (>90%).[5,16] In settings where *Brucella* is endemic, serological tests, including the standard tube agglutination, Rose Bengal, and Coombs anti-*Brucella* tests, are essential because culture sensitivity is limited for slow-growing *Brucella* spp.[9] In our practice, the primary route to microbiological diagnosis remains intraoperative tissue sampling performed during debridement. A further contributor specific to tertiary academic practice in Türkiye is the referral chain itself: most patients reach a neurosurgical centre only after a period of empirical oral antibiotic therapy initiated in primary or secondary care for non-specific back pain, well before magnetic resonance imaging and microbiological sampling are pursued. This pretreatment effect is plausibly an important driver of the high culture-negative proportion observed in surgically referred cohorts and contrasts with the higher microbiological yield reported in series with early infectious diseases involvement and prompt biopsy.

The dominant role of abscesses in shaping treatment is consistent with IDSA recommendations[5] and recent neurosurgical evidence, as abscess presence was associated with both the decision to operate (Fisher's exact p=0.009) and the duration of intravenous antibiotic therapy (p=0.008). While Joerger et al.[4] demonstrated that epidural abscess is a significant prognostic factor for early mortality associated with spondylodiscitis, Motov et al.[17] reported that the presence of complications such as abscesses determines the indication for radical surgical debridement in cervical spondylodiscitis. The newly proposed Spinal Infection Treatment Evaluation (SITE) score similarly incorporates abscesses and neurological deficits as central variables in the surgical decision-making process.[18] Our finding that 34 of 35 patients (97.1%) with an abscess underwent surgery supports the current view that an epidural or paravertebral abscess is one of the strongest indications for surgical intervention, although an abscess alone is not universally regarded as an absolute indication in international guidelines.

Regarding the duration of antibiotic therapy, the median duration of 36 days in our cohort was below the recommended six-week threshold for pyogenic disease when the entire population was considered.[5,11] However, stratification by granulomatous and pyogenic disease revealed that pyogenic SD required approximately twice as long a course of intravenous antibiotic treatment as granulomatous SD (median 43 days vs. 21 days; p=0.021). This reflects only the intravenous phase: brucellosis and tuberculosis regimens are administered predominantly orally following a short parenteral induction phase.[9,15] A landmark study by Bernard et al.[11] demonstrated that a six-week course of antibiotic treatment for pyogenic vertebral osteomyelitis was no less effective than a twelve-week coursealthough most of these patients received only medical treatment and the prevalence of abscesses was low. Our data indicate that, in the pyogenic subgroup of this sur-

gical cohort with a high abscess burden, intravenous treatment durations approached or exceeded six weeks, in line with observations from other neurosurgical centres.[13,17] This is facilitated by the high oral bioavailability of doxycycline, rifampicin, and isoniazid and reduces hospital stay and catheter-related risks, whereas pyogenic infections, particularly methicillin-resistant staphylococcal infections requiring glycopeptide therapy, may depend on prolonged parenteral administration to achieve adequate bone and disc penetration. The seven conservatively managed patients (three with brucellosis and four culture-negative) had a short median intravenous course of 8 days, consistent with early oral step-down rather than inadequate antimicrobial exposure and with the referral logistics whereby non-operative patients continue oral management under Infectious Diseases.

This study has several limitations. First, the single-centre retrospective design and limited sample size restrict generalisability and precluded multivariable adjustment, particularly for the small conservatively managed subgroup (n=7); the association between surgical intervention and prolonged antibiotic therapy should therefore be interpreted in the context of selection by indication, as surgically treated patients generally had more extensive infection at baseline. Second, the high culture-negative rate limited the granulomatous-versus-pyogenic comparison to 28 patients. Third, long-term data on recurrence, mortality, and functional outcomes were not systematically available. Fourth, the absence of serial inflammatory markers (CRP, ESR) and imaging follow-up precluded biomarker-guided assessment of treatment duration. Finally, as a neurosurgical series from a *Brucella*-endemic, livestock-farming region of Türkiye, uncomplicated brucellar spondylodiscitis without neurological deficit, instability, or a significant abscess, which is typically managed conservatively by Infectious Diseases, is under-represented; cases referred to neurosurgery preferentially involve neurological compromise, instability, or substantial epidural or paravertebral collections, which increases the observed proportion of complicated, abscess-associated disease and limits extrapolation of our microbiological distribution to the general regional population.

Practically, these findings argue against pyrexia as a triage criterion for suspected spondylodiscitis: persistent unexplained back pain with elevated inflammatory markers should suffice to trigger urgent MRI. They also support extending brucellar serology and tuberculosis-directed testing to every patient in endemic provinces, irrespective of occupational exposure history, rather than reserving such testing for those with a suggestive exposure or contact history. Future multicentre studies combining neurosurgical and infectious diseases cohorts are needed to validate abscess-centred risk-stratification instruments, such as the SITE score, in brucellosis-endemic settings.

CONCLUSION

In a neurosurgical cohort from western Türkiye, where *Brucella* is endemic, granulomatous pathogens predominated over *S. aureus*, and nearly half of all cases remained culture-negative despite modern diagnostic methods. The presence of an abscess emerged as the primary determinant of both surgical treatment and prolonged intravenous antimicrobial therapy. Stratified analysis revealed that pyogenic SD required approximately twice as long a course of intravenous antibiotic treatment as granulomatous SD. These findings highlight the importance of diagnostic algorithms adapted to local conditions, including routine serological testing for brucellosis, and a pragmatic, abscess-focused approach to surgical and antimicrobial decision-making. Multicentre prospective studies are needed to further clarify treatment-duration criteria and prognostic factors in Türkiye.

Disclosures

Ethics Committee Approval: The study was approved by the Balıkesir University Faculty of Medicine Health Research Clinical Research Ethics Committee (No: 2026/9-7, Date: 05/05/2026).

Informed Consent: The requirement for informed consent was waived by the ethics committee due to the retrospective nature of the study.

Conflict of Interest Statement: The authors declare no conflicts of interest related to this work.

Funding: The authors received no specific funding for this work.

Use of AI for Writing Assistance: A large language model (Anthropic Claude, accessed via web interface, 2026) was used to support English-language editing, manuscript formatting, table construction, and figure generation. The AI tool did not generate study data, perform clinical interpretation, or determine scientific conclusions. All content was reviewed and verified by the authors, who accept full responsibility for the integrity, accuracy, and originality of the manuscript, including ensuring appropriate attribution of all quoted material.

Author Contributions: Concept – S.K.; Design – S.K.; Supervision – S.K., A.C.K., M.F.S., M.A.T., R.İ.S., E.F., H.U.A.; Funding – S.K., A.C.K., M.F.S., M.A.T., R.İ.S., E.F., H.U.A.; Materials – S.K., A.C.K., M.F.S., M.A.T., R.İ.S., E.F., H.U.A.; Data collection and/or processing – S.K., A.C.K., M.F.S., E.F., H.U.A.; Data analysis and/or interpretation – M.A.T., R.İ.S., S.K., A.C.K.; Literature search – S.K., A.C.K., M.F.S., M.A.T., R.İ.S., E.F., H.U.A.; Writing – S.K., A.C.K.; Critical review – S.K., A.C.K., M.F.S., M.A.T., R.İ.S., E.F., H.U.A.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Lacasse M, Derolez S, Bonnet E, Amelot A, Bouyer B, Carlier R, et al. 2022 SPILF - Clinical Practice guidelines for the diagnosis and treatment of disco-vertebral infection in adults. Infect Dis Now 2023;53:104647. [CrossRef]
2. Crombé A, Fadli D, Clinca R, Reverchon G, Cevolani L, Girolami M, et al. Imaging of spondylodiscitis: a comprehensive updated review-multimodality imaging findings, differential diagnosis, and specific microorganisms detection. microorganisms. 2024;12:93. [CrossRef]
3. Sobottke R, Seifert H, Fätkenheuer G, Schmidt M, Gossmann A, Eysel P. Current diagnosis and treatment of spondylodiscitis. Dtsch Arztebl Int 2008;105:181–7. [CrossRef]
4. Joerger AK, Albrecht C, Lange N, Meyer B, Wostrack M. In-hospital mortality from spondylodiscitis: insights from a single-center retrospective study. J Clin Med 2023;12:7228. [CrossRef]
5. Berbari EF, Kanj SS, Kowalski TJ, Darouiche RO, Widmer AF, Schmitt SK, et al. 2015 Infectious Diseases Society of America (IDSA) clinical practice guidelines for the diagnosis and treatment of native vertebral osteomyelitis in adults. Clin Infect Dis 2015;61–46. [CrossRef]
6. Akcam FZ, Kaya O, Ceylan T. Comment on: Spondylodiscitis: update on diagnosis and management. J Antimicrob Chemother 2011;66:1199–200. [CrossRef]
7. Bal E, Topcu E, Yuluğkural Z. Risk factors and clinical characteristics of spondylodiscitis: a comparative study. Int Orthop 2025;49:1461–70. [CrossRef]
8. Karadağ MK, Alay H, Çankaya BY. Spondylodiscitis as osteoarticular involvement in brucellosis. Rev Soc Bras Med Trop 2023;56. [CrossRef]
9. Esmaeilnejad-Ganji SM, Esmaeilnejad-Ganji SMR. Osteoarticular manifestations of human brucellosis: A review. World J Orthop 2019;10:54–62. [CrossRef]
10. Kaya S, Kaya S, Kavak S, Comoglu S. A disease that is difficult to diagnose and treat: evaluation of 343 spondylodiscitis cases. J Int Med Res 2021;49:3000605211060197. [CrossRef]
11. Bernard L, Dinh A, Ghout I, Simo D, Zeller V, Issartel B, et al. Antibiotic treatment for 6 weeks versus 12 weeks in patients with pyogenic vertebral osteomyelitis: an open-label, non-inferiority, randomised, controlled trial. Lancet 2015;385:875–82. [CrossRef]
12. Lima D, Lopes N, Pereira AL, Rodrigues D, Amaral-Silva M, Marques E. Diagnosis and treatment of spondylodiscitis: findings from a five-year single-centre study. Cureus 2024;16(11). [CrossRef]
13. Tschugg A, Lener S, Hartmann S, Rietzler A, Neururer S, Thomé C. Primary acquired spondylodiscitis has a more severe course than post-spinal surgery spondylodiscitis: a single-centre retrospective study of 159 cases. Neurosurg Rev 2018;41:141–7. [CrossRef]
14. Conti J, Geremia N, Di Bella S, Zadra F, Zerbato V, Babich S, et al. Infectious spondylodiscitis: epidemiology, diagnosis, microbiological findings, clinical features and outcomes in a 14-year retrospective study. Eur J Clin Microbiol Infect Dis 2025;44:3067–78. [CrossRef]
15. Ariza J, Bosilkovski M, Cascio A, Colmenero JD, Corbel MJ, Falagas ME, et al. Perspectives on the treatment of brucellosis in the 21st century: the Ioannina recommendations. PLoS Med 2007;4. [CrossRef]
16. Yu D, Kang Y, Lu W, Chen B. Progress in diagnosis and treatment of primary spondylodiscitis: a systematic literature review. EFORT Open Rev 2025;10:815–28. [CrossRef]
17. Motov S, Stemmer B, Krauss P, Bonk MN, Wolfert C, Steininger K, et al. Clinical and surgical outcomes in patients with cervical spondylodiscitis-a single-centre retrospective case series of 24 patients. Front Surg 2024;11:1292977. [CrossRef]
18. Campisi BM, Costanzo R, Lo Bue E, Rubino R, Panico F, Cipollina GP, et al. Evaluation of the decision-making process for treatment in spontaneous spondylodiscitis: a retrospective study involving two neurosurgical centres. World Neurosurg 2025;199:123994. [CrossRef]

DOI: 10.14744/cm.2026.78055
Compreh Med 2026;18(Suppl. 1):112-118

RESEARCH ARTICLE

Admission Neutrophil-to-Lymphocyte Ratio and Hospital Stay After Surgery for Chronic Subdural Hematoma Outcomes

Fatih Çalış[1], Mohammed Aladdam[2], Mehmet Ali Kahraman[1], Said Abumarasa[3], Ahmet Faruk Özdemir[1], Semih Naci Balak[1]

[1]Department of Neurosurgery, İstanbul Medeniyet University, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, İstanbul, Türkiye
[2]Department of Neurosurgery, Academic Hospital, İstanbul, Türkiye
[3]Department of Radiology, Medipol Health Group, Medipol Bahçelievler Hospital, İstanbul, Türkiye

ABSTRACT

Objective: Chronic subdural hematoma (CSDH) is increasingly prevalent among older adults and is characterized by inflammation-driven pathophysiology. The objective of this study was to evaluate whether the admission NLR predicts postoperative length of stay (PoLOS) in surgically treated patients with CSDH and to examine its association with early quantitative radiological hematoma resolution.

Materials and Methods: This retrospective, single-center study included 111 adult patients who underwent burr-hole evacuation with subdural drainage for CSDH. Admission NLR was calculated from peripheral blood counts. The primary outcome was postoperative LOS ($\log_{10}$-transformed for regression analysis). The secondary analysis assessed the percentage reduction in hematoma thickness on early postoperative and approximately two-week CT scans. Univariable and multivariable linear regression models were performed with adjustment for clinically relevant covariates.

Results: Higher admission NLR was associated with longer postoperative LOS in the univariable analysis (β=0.009, p=0.028). After adjustment for age and anticoagulant use, the association was attenuated and did not reach statistical significance (β=0.007, p=0.074). Age remained independently associated with postoperative LOS (β=0.003 per year, p=0.024). Admission NLR was not correlated with early postoperative hematoma reduction (ρ=0.008, p=0.933) or two-week hematoma reduction (ρ=-0.100, p=0.295).

Conclusion: Admission NLR was associated with longer postoperative LOS in the univariable analysis but not after adjustment, warranting further investigation in larger cohorts.

Keywords: Chronic subdural hematoma, postoperative hospital length of stay, neutrophil-to-lymphocyte ratio, prognostic biomarker, surgical outcomes, systemic inflammation

How to cite this article: Çalış F, Aladdam M, Kahraman MA, Abumarasa S, Özdemir AF, Balak SN. Admission Neutrophil-to-Lymphocyte Ratio and Hospital Stay After Surgery for Chronic Subdural Hematoma Outcomes. Compreh Med 2026;18(Suppl. 1):112-118.

INTRODUCTION

Chronic subdural hematoma (CSDH) is a common neurosurgical condition, particularly in elderly patients, and its incidence continues to rise with population aging and the increasing use of antithrombotic therapy. Although surgical evacuation via burr-hole drainage is generally effective, postoperative radiological evolution varies substantially among patients, with some demonstrating rapid resolution while others exhibit persistent residual collections or delayed clearance. This heterogeneity suggests that factors beyond mechanical evacuation alone contribute to hematoma persistence and resolution.[1–3]

Accumulating evidence indicates that CSDH is not a passive accumulation of blood but rather a chronic inflammatory disease. The outer hematoma membrane is characterized by inflammatory cell infiltration, angiogenesis, and recur-

Address for Correspondence: Fatih Çalış, Department of Neurosurgery, İstanbul Medeniyet University, Göztepe Prof. Dr. Süleyman Yalçın City Hospital, İstanbul, Türkiye
E-mail: fthcls@hotmail.com **ORCID ID:** 0000-0003-2637-1152

Received date: 01.07.2026
Revised date: 15.09.2026
Accepted date: 17.09.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

rent microhemorrhage, driven by a complex cytokine and growth-factor milieu. These biological processes underpin both hematoma expansion and recurrence and may also influence postoperative resorption dynamics.[2,4,5]

Systemic inflammatory markers have therefore attracted interest as potential predictors of CSDH behavior. Among these, the neutrophil-to-lymphocyte ratio (NLR) is a readily available and inexpensive biomarker reflecting the balance between innate inflammatory activation and adaptive immune regulation. Elevated NLR has been associated with adverse outcomes across a wide range of neurological and neurosurgical conditions, including intracerebral hemorrhage, traumatic brain injury, and brain tumors.[6–8]

In chronic subdural hematoma (CSDH), several studies have associated the neutrophil-to-lymphocyte ratio (NLR) with clinical outcomes, primarily focusing on hematoma recurrence or disease progression. Elevated preoperative or postoperative NLR values, as well as postoperative increases in NLR, have been linked to higher recurrence rates after surgical evacuation and to progression during conservative management.[9–11] Recently, Kartal et al.[12] examined peripheral inflammatory indices in relation to hospital length of stay in surgically treated patients with CSDH younger than 65 years, excluding patients with coagulopathy, anticoagulant or antiplatelet drug use, and bilateral hematomas, and reported an association between perioperative changes in NLR and hospitalization duration. However, CSDH predominantly affects older adults, among whom anticoagulant use is common; importantly, these patients were included in our cohort. Therefore, the relationship between NLR measured at admission and postoperative length of stay (PoLOS) in a broader, predominantly older CSDH population remains incompletely characterized. We specifically evaluated postoperative LOS because inclusion of the preoperative period could introduce additional nonclinical and treatment-related determinants of hospitalization duration, such as surgical delay related to anticoagulant use or waiting for operative scheduling in neurologically stable patients.

Postoperative hospital length of stay (PoLOS) is a clinically meaningful and resource-sensitive outcome that reflects the overall postoperative course, encompassing neurological recovery, medical complications, and readiness for discharge. Although PoLOS is multifactorial, systemic inflammatory burden may plausibly contribute to delayed recovery and prolonged hospitalization in patients undergoing surgical treatment for CSDH.

In contrast, quantitative radiological evolution—including the percentage of early postoperative hematoma resolution and short-term reduction at approximately two weeks—has received comparatively little attention in relation to systemic inflammation. Early radiological change may provide supportive biological insight into postoperative hematoma dynamics, as it reflects the balance between residual inflammatory activity and hematoma resorption. Evaluating radiological resolution alongside postoperative LOS may therefore help contextualize the clinical impact of inflammation in the early postoperative period without relying solely on dichotomous outcomes such as recurrence.

Accordingly, the objective of this study was to evaluate whether the neutrophil-to-lymphocyte ratio (NLR) predicts PoLOS in patients undergoing surgical treatment for chronic subdural hematoma. As a secondary analysis, we examined the association between NLR and early quantitative radiological hematoma resolution, expressed as the percentage of hematoma reduction in the immediate postoperative period and at approximately two-week follow-up.

MATERIALS and METHODS

Study Design and Ethical Approval

This retrospective, single-center observational study was performed at a tertiary referral neurosurgical center. Consecutive adult patients who underwent surgical treatment for chronic subdural hematoma (CSDH) between March 2018 and March 2025 were identified from institutional surgical and radiological databases. The study protocol was reviewed and approved by the Göztepe Prof. Dr. Süleyman Yalçın City Hospital Ethics Review Board (Approval No: 2025/0369; Date: 18/12/2025) and was conducted in accordance with the principles of the Declaration of Helsinki.

Patient Population and Eligibility Criteria

Patients were eligible for inclusion if they were adults with radiologically confirmed CSDH who underwent burr-hole evacuation with placement of a subdural drain during the study period. Exclusion criteria included incomplete clinical or imaging data, absence of drain placement, lack of postoperative CT imaging, primary craniotomy or other nonstandard surgical approaches, and the presence of concurrent acute medical conditions that could substantially influence neurological status or perioperative recovery, including but not limited to significant electrolyte disturbances, acute exacerbation of chronic obstructive pulmonary disease, or acute cardiovascular events such as myocardial infarction.

In cases of bilateral CSDH treated during the same surgical session, both hematomas were managed simultaneously and analyzed as a single clinical case.

Surgical Technique

All procedures were performed using a standard burr-hole craniostomy technique under either general or local anesthesia, according to surgeon and anesthetic preference. Following dural opening and irrigation of the subdural space, a subdural drain was inserted through the burr hole, with the catheter tip positioned within the subdural compartment. All operations were performed by using uniform institutional surgical protocols.

Data Collection

Clinical, laboratory, and radiological data were retrospectively extracted from electronic medical records and the institutional picture archiving and communication system (PACS). Baseline demographic and clinical variables included patient age and sex, history of hypertension, diabetes mellitus, and prior stroke, use of anticoagulant medication at presentation, composite score, and documented head trauma, with the interval between trauma and hospital admission recorded when applicable. Neurological status at admission was assessed using the Glasgow Coma Scale (GCS), and the presence of focal neurological deficits was recorded as a dichotomous variable. A focal deficit was defined as muscle strength graded ≤4/5 on the Medical Research Council (MRC) scale.

Radiological Assessment

Radiological evaluation was based on noncontrast cranial CT scans obtained preoperatively, in the early postoperative period, and at approximately 14 days postoperatively. Hematoma thickness was measured on coronal images as the maximum subdural width (mm) on the operated side. In patients with bilateral hematomas, measurements from both sides were obtained separately and summed to yield the total hematoma thickness. Midline shift was measured in millimeters at the level of the foramen of Monro on preoperative CT as an indicator of mass effect. Hematoma laterality (right, left, or bilateral) and the presence of internal membranes or septations within the hematoma cavity were also recorded.

Outcome Measures

The primary clinical outcome was postoperative hospital length of stay (PoLOS), defined as the number of days from surgery to hospital discharge. As a supportive radiological assessment, changes in subdural hematoma thickness were quantified as the percentage reduction in thickness across the predefined imaging time points (preoperative, early postoperative, and 14-day postoperative). Secondary outcomes included intensive care unit (ICU) admission (yes/no) and hematoma recurrence, defined as symptomatic reaccumulation requiring repeat surgical evacuation during follow-up. Postoperative neurological status was assessed using the GCS obtained either in the early postoperative period or at discharge.

Baseline disease severity was characterized using preoperative midline shift and Markwalder grading, which were used descriptively to characterize clinical and radiological severity profiles in subsequent analyses.

Ethical Statement

This study was conducted in accordance with the ethical standards of the institutional research committee and the 1964 Declaration of Helsinki and its later amendments. The study protocol was reviewed and approved by the Göztepe Prof. Dr. Süleyman Yalçın City Hospital Ethics Review Board (Approval No: 2025/0369; Date: 18/12/2025).

Statistical Analysis

Continuous variables are presented as mean±standard deviation (SD) or median with interquartile range (IQR), depending on their distribution, and categorical variables are summarized as counts and percentages. Normality of continuous variables was assessed using the Kolmogorov–Smirnov and Shapiro–Wilk tests. Baseline demographic, clinical, and radiological characteristics—including preoperative midline shift and Markwalder grade as indicators of disease severity—were compared between groups using independent-samples t-tests or Mann–Whitney U tests for continuous variables and χ^2 or Fisher's exact tests for categorical variables, as appropriate.

The primary outcome was postoperative hospital length of stay (PoLOS), defined as the number of days from surgery to discharge. Because LOS showed a right-skewed distribution, a $\log_{10}$ transformation was applied prior to regression analysis. The association between the neutrophil-to-lymphocyte ratio (NLR) and PoLOS was evaluated using multiple linear regression, with $\log_{10}$-transformed LOS as the dependent variable. Potential covariates were initially examined individually. Given the exploratory nature of the study and the relatively limited sample size, a multivariable model including age, anticoagulant use, and admission NLR was used to assess whether the observed NLR–PoLOS association persisted after adjustment for selected clinically relevant factors.

Binary outcomes, including ICU admission and hematoma recurrence requiring reoperation, were summarized descriptively and compared using χ^2 or Fisher's exact tests, depending on expected cell counts. Given the limited number of events, multivariable logistic regression was not performed for these outcomes. All statistical analyses were conducted using IBM SPSS Statistics version 26 (IBM Corp., Armonk, NY, USA), and a two-sided p value <0.05 was considered statistically significant.

RESULTS

A total of 111 patients who underwent surgical treatment for chronic subdural hematoma were included in the analysis (Table 1). The median age was 74 years (interquartile range [IQR], 19 years), and 71 patients (64.0%) were male. Hematoma laterality was right-sided in 42 patients (37.8%), left-sided in 48 patients (43.2%), and bilateral in 21 patients (18.9%). A history of head trauma recalled by the patient or relatives was reported in 62 patients (55.9%).

Hypertension, diabetes mellitus, and prior cerebrovascular disease were present in 66 (59.5%), 24 (21.6%), and 15 (13.5%) patients, respectively. Based on these conditions, the composite comorbidity score was 0 in 41 patients (36.9%), 1 in 39 (35.1%), 2 in 27 (24.3%), and 3 in 4 (3.6%). Thirty patients (27.0%) were receiving anticoagulant therapy at admission.

Radiologically, the mean total hematoma thickness—calculated as the sum of bilateral thicknesses in patients with bilateral hematomas—was 25.9±9.5 mm (range, 10.4–55.5 mm), and the median midline shift was 5.8 mm (IQR, 9.6 mm; range, 0–19.0 mm). At presentation, the median Glasgow Coma Scale score was 15 (IQR, 2; range, 7–15), and the median Markwalder grade was 1 (IQR, 1; range, 0–4).

Regarding inflammatory markers at admission, the median neutrophil count was 5.21×10^9/L (IQR, 3.40×10^9/L), the median lymphocyte count was 1.60×10^9/L (IQR, 1.10×10^9/L), and the median neutrophil-to-lymphocyte ratio (NLR) was 3.46 (IQR, 4.29), demonstrating a right-skewed distribution.

In univariable linear regression analyses with $\log_{10}$-transformed PoLOS as the dependent variable, a higher preoperative neutrophil-to-lymphocyte ratio was significantly associated with longer hospitalization (β=0.009, p=0.028), corresponding to an approximately 2% increase in PoLOS per unit increase in NLR. Increasing age was also significantly associated with prolonged PoLOS (β=0.004 per year, p=0.002), equating to an approximately 10% increase per decade. Patients receiving anticoagulant therapy at presentation experienced significantly longer PoLOS than non-anticoagulated patients (β=0.153, p=0.006). A higher Markwalder grade was likewise associated with increased PoLOS (β=0.066, p=0.021), reflecting greater baseline clinical severity. In contrast, sex, overall comorbidity burden, total hematoma thickness, and preoperative midline shift were not significantly associated with PoLOS in univariable analyses. Postoperative neutrophil-to-lymphocyte ratio was also not associated with PoLOS.

Table 1. Baseline demographic, clinical, and radiological characteristics of patients undergoing surgical treatment for chronic subdural hematoma (n=111)

Variable	Value (n=111)
Age (years)	69.2±18.5 [74±20]
Male sex, n (%)	71 (64.0)
Anticoagulant use, n (%)	30 (27.0)
Bilateral hematoma, n (%)	24 (21.6)
Neurological deficit at presentation, n (%)	37 (33.3)
Radiological membranes present, n (%)	65 (58.6)
Hypertension, n (%)	66 (59.5)
Diabetes mellitus, n (%)	24 (21.6)
History of CVA, n (%)	15 (13.5)
History of trauma, n (%)	62 (55.9)
Total hematoma load (mm)	25.6±9.2 [25.2±12.5]
Midline shift (mm)	6.1±4.8 [5.9±10.1]
Markwalder grade, n (%)	
Grade 0	24 (21.8)
Grade 1	44 (40.0)
Grade 2	36 (32.7)
Grade 3	5 (4.5)
Grade 4	1 (0.9)
Preoperative NLR	5.79±6.09 [3.45±4.32]
Early postoperative thickness (mm)	7.2±3.5 [6.45±4.25]
Second-week postoperative thickness (mm)	6.5±4.4 [5.50±5.75]
Postoperative NLR	7.82±7.87 [5.35±6.77]
Postoperative GCS	14.5±1.4 [15±0]
PoLOS (days)	8.45±8.91 [6±5]
Log10 PoLOS	0.82±0.26 [0.78±0.35]
Recurrence requiring surgery, n (%)	4 (3.6)

Continuous variables are presented as mean ± standard deviation [median± interquartile range]. Categorical variables are expressed as number (n) and percentage (%). Percentages were calculated using valid cases. Markwalder grade was available for 110 patients. NLR: Neutrophil-to-lymphocyte ratio; GCS: Glasgow coma scale; CVA: Cerebrovascular accident; PoLOS: Postoperative hospital length of stay

A multivariable model including admission NLR, age, and anticoagulant use was subsequently evaluated.

In multivariable linear regression analysis adjusting for age and anticoagulant use, preoperative neutrophil-to-lymphocyte ratio demonstrated a trend toward an association with PoLOS (β=0.007, p=0.074), with higher values corresponding to longer hospitalization. Increasing age remained independently associated with prolonged PoLOS (β=0.003 per year, p=0.024), while anticoagulant use showed a borderline association with longer hospital stay (β=0.107, p=0.056).

Regarding radiological resolution, quantitative analysis demonstrated a substantial reduction in total subdural hematoma thickness following surgical evacuation. The mean early postoperative reduction in hematoma thickness was 72.2%±8.6% (median, 72.4%; IQR, 9.6%). At approximately two-week follow-up, the mean reduction increased to 75.5%±14.9% (median, 76.9%; IQR, 16.4%), indicating continued hematoma resolution over time.

Spearman correlation analysis demonstrated no significant association between preoperative neutrophil-to-lymphocyte ratio and either early postoperative radiological hematoma reduction (ρ=0.008, p=0.933) or the magnitude of hematoma reduction at approximately two-week follow-up (ρ=-0.100, p=0.295).

These findings indicate that while baseline systemic inflammation is associated with prolonged hospitalization, it does not appear to influence the early radiological resolution of chronic subdural hematoma.

DISCUSSION

In this study, higher admission NLR was significantly associated with longer PoLOS in univariable analysis. After adjustment for age and anticoagulant use, the magnitude and direction of the association remained broadly consistent, although it no longer reached conventional statistical significance. In contrast, admission NLR was not associated with early postoperative or short-term radiological hematoma resolution. These findings do not establish NLR as an independent predictor of postoperative LOS but suggest that admission NLR may capture aspects of inflammatory and physiological vulnerability that overlap with other clinical determinants of postoperative recovery.

The NLR was originally proposed by Zahorec (2001) as a rapid, simple marker of systemic inflammatory stress in critically ill patients.[6] It has since evolved into a widely used surrogate of disease severity across acute and chronic conditions. Early "severity" studies subsequently showed that higher NLR is associated with worse clinical phenotypes and outcomes in high-burden fields such as cardiovascular disease, including associations with impaired coronary perfusion and higher in-hospital and long-term mortality after STEMI treated with PCI, and oncology, where NLR is frequently incorporated into prognostic frameworks and linked to poorer survival.[13] Importantly, the literature has also established a relationship between NLR and resource utilization, particularly prolonged hospitalization, in several settings where inflammatory dysregulation parallels clinical complexity. In acute ischemic stroke, for example, Zhao et al.[14] demonstrated that admission NLR independently predicted longer length of stay and higher acute hospital costs, complementing broader evidence connecting inflammation-related markers with stroke severity and early deterioration. Similar LOS associations have been reported in other acute neurological syndromes, such as status epilepticus, where admission NLR predicted length of hospitalization and ICU need. The concept has also recently extended into psychiatry, where large multicenter analyses have shown that elevated NLR and related inflammatory ratios are associated with longer inpatient stays independent of diagnosis and site.[15,16]

Taken together, these data reinforce the concept that NLR reflects not only biological disease activity but also clinical trajectory and healthcare resource utilization. In this context, our findings are consistent with the broader literature: higher preoperative NLR was associated with prolonged postoperative length of stay (PoLOS), suggesting that even in chronic subdural hematoma—a condition not traditionally conceptualized as primarily inflammatory—systemic inflammatory status may influence recovery dynamics and the inpatient course. Our findings are therefore consistent with a potential relationship between systemic inflammatory status and postoperative clinical trajectory. However, because the association between admission NLR and postoperative LOS did not remain statistically significant after adjustment, the present results should be considered exploratory and do not establish NLR as an independent prognostic biomarker for hospitalization duration.

An elevated NLR likely reflects increased systemic inflammation and physiological stress, which may slow recovery and delay discharge even after successful hematoma evacuation. In contrast, early radiological reduction in hematoma thickness is primarily determined by surgical evacuation and brain expansion and may therefore be less influenced by systemic inflammatory status in the short term.

The earliest study to evaluate NLR in the setting of chronic subdural hematoma appears to be the report by Zhang et al.,[17] who examined peripheral blood inflammatory markers in a large CSDH cohort and found higher NLR and PLR values in patients with CSDH than in healthy controls. Importantly, their work primarily supported the concept of CSDH as an inflammatory condition and highlighted the need for subsequent studies to clarify the prognostic role of these indices. Building on this foundation, later studies shifted from case–control comparisons toward outcome prediction, most commonly focusing on recurrence risk and postoperative inflammatory dynamics.

Dobran et al.[9] specifically evaluated preoperative NLR and reported that higher values were associated with an increased risk of postoperative recurrence, proposing NLR as a simple preoperative risk marker. In contrast, de Oliveira et al.[10] emphasized postoperative inflammatory dynamics, showing that patients who experienced recurrence had higher postoperative neutrophil counts and NLR values and that the postoperative-to-preoperative NLR ratio was associated with recurrence, whereas preoperative laboratory parameters were not predictive in their cohort.

In 2025, Yang et al.[11] broadened the question to hematoma progression rather than recurrence after surgery and found that NLR, alongside coagulation and comorbidity factors such as INR, hypertension, diabetes, and WBC, was independently associated with progression in their retrospective analysis. Taken together, these studies consistently support a relationship between systemic inflammatory markers and "harder" CSDH behaviors such as recurrence or progression, whereas our results add nuance by suggesting that admission NLR may relate more to early clinical recovery and resource utilization, such as PoLOS, than to short-term, mechanically driven radiological thickness reduction.

More recently, Kartal et al.[12] examined peripheral inflammatory indices in relation to clinical outcomes, including hospital length of stay, in 53 surgically treated patients with CSDH younger than 65 years. Notably, the preoperative N/L ratio itself was not significantly correlated with hospital length of stay ($r=-0.128$, $p=0.360$), whereas the perioperative change in the N/L ratio showed a significant positive correlation with hospitalization duration ($r=0.562$, $p=0.001$). Their findings therefore suggest that perioperative inflammatory dynamics may be related to hospitalization duration. The present study addresses a related but distinct question by evaluating NLR measured at admission in a larger, predominantly older CSDH cohort and examining its association with PoLOS using regression analysis with adjustment for relevant clinical covariates. In addition, we evaluated admission NLR in relation to quantitative early postoperative hematoma resolution. These differences extend the available evidence by examining whether inflammatory status already present at admission is associated with the subsequent postoperative course.

Limitations

This study has several limitations. First, its retrospective and single-center design limits the ability to establish definitive associations and may reduce generalizability to other institutions. Second, PoLOS is influenced by multiple factors beyond medical recovery, including institutional practices and discharge planning, particularly for older adults, which could not be fully accounted for.

Third, the neutrophil-to-lymphocyte ratio is a nonspecific marker of systemic inflammation and may be affected by comorbid conditions or physiological stress unrelated to chronic subdural hematoma. Other inflammatory markers were not evaluated, and postoperative changes in NLR over time were not analyzed. Radiological resolution was assessed only in the early postoperative and short-term follow-up periods; longer-term imaging and functional outcomes were not examined.

Fourth, although the Markwalder grade was associated with postoperative LOS in univariable analysis, it was not included in the final multivariable model. Residual confounding by baseline neurological severity therefore cannot be excluded, and the adjusted association between NLR and LOS should be interpreted cautiously.

Finally, radiological evolution was assessed using maximal hematoma thickness rather than volumetric measurements; therefore, particularly in bilateral or irregularly shaped hematomas, this approach may not fully capture changes in total hematoma volume.

CONCLUSION

In patients undergoing surgical treatment for CSDH, higher admission NLR was associated with longer postoperative LOS in univariable analysis, but this association did not remain statistically significant after adjustment for age and anticoagulant use. Admission NLR was not associated with early radiological hematoma resolution. These findings suggest a possible relationship between admission inflammatory status and postoperative recovery, but the independent prognostic value of NLR remains uncertain and requires confirmation in larger studies.

Disclosures

Ethics Committee Approval: The study was approved by the Göztepe Prof. Dr. Süleyman Yalçın City Hospital Ethics Committee (No: 2025/0369, Date: 18/12/2025).

Informed Consent: Informed consent was waived due to study design.

Conflict of Interest Statement: The authors have no relevant financial or non-financial interests to disclose.

Funding: This research received no external funding.

Use of AI for Writing Assistance: ChatGPT 5.0 was used for language editing and drafting support.

Author Contributions: Concept – F.Ç., M.A., M.A.K.; Design – F.Ç., M.A., M.A.K., S.N.B.; Supervision – F.Ç., S.N.B.; Materials – M.A.K., S.A., A.F.Ö.; Data collection and/or processing – M.A.K.; Data analysis and/or interpretation – M.A., F.Ç., M.A.K., S.A.; Writing – F.Ç., A.F.Ö.; Critical review – F.Ç., S.N.B.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Santarius T, Kirkpatrick PJ, Ganesan D, Chia HL, Jalloh I, Smielewski P, et al. Use of drains versus no drains after burr-hole evacuation of chronic subdural haematoma: a randomised controlled trial. Lancet 2009;374:1067–73. [CrossRef]
2. Edlmann E, Giorgi-Coll S, Whitfield PC, Carpenter KLH, Hutchinson PJ. Pathophysiology of chronic subdural haematoma: inflammation, angiogenesis and implications for pharmacotherapy. J Neuroinflammation 2017;14:108. [CrossRef]
3. Holl DC, Volovici V, Dirven CMF, Peul WC, van Kooten F, Jellema K, et al. Pathophysiology and nonsurgical treatment of chronic subdural hematoma: from past to present to future. World Neurosurg 2018;116:402–11.e2. [CrossRef]
4. Weigel R, Schilling L, Krauss JK. The pathophysiology of chronic subdural hematoma revisited: emphasis on aging processes as key factor. Geroscience 2022;44:1353–71. [CrossRef]
5. Stanišić M, Hald J, Rasmussen IA, Pripp AH, Ivanović J, Kolstad F, et al. Volume and densities of chronic subdural haematoma obtained from CT imaging as predictors of postoperative recurrence: a prospective study of 107 operated patients. Acta Neurochir (Wien) 2013;155:323–33; discussion 333. [CrossRef]
6. Zahorec R. Ratio of neutrophil to lymphocyte counts--rapid and simple parameter of systemic inflammation and stress in critically ill. Bratisl Lek Listy 2001;102:5–14. [CrossRef]
7. Tokgoz S, Kayrak M, Akpinar Z, Seyithanoğlu A, Güney F, Yürüten B. Neutrophil lymphocyte ratio as a predictor of stroke. J Stroke Cerebrovasc Dis 2013;22:1169–74. [CrossRef]
8. Lattanzi S, Cagnetti C, Provinciali L, Silvestrini M. Neutrophil-to-lymphocyte ratio predicts the outcome of acute intracerebral hemorrhage. Stroke 2016;47:1654–7. [CrossRef]
9. Dobran M, Carrassi E, Di Rienzo A, Lattanzi S, Capece M, Aiudi D, Iacoangeli M. Preoperative neutrophil to lymphocyte ratio as a predictor of chronic subdural hematoma recurrence. Interdisciplinary Neurosurgery 2022;28:101436. [CrossRef]
10. de Oliveira AJM, Solla DJF, de Oliveira KF, Amaral BS, Andrade AF, Kolias AG, et al. Postoperative neutrophil-to-lymphocyte ratio variation is associated with chronic subdural hematoma recurrence. Neurol Sci 2022;43:427–34. [CrossRef]
11. Yang W, Du Z, Chen Q, Yuan Q, Fu P, Fang J, et al. Factors associated with progression and sex-specific patterns in chronic subdural hematoma: a single-centre retrospective study. Front Neurol 2025;16:1674327. [CrossRef]
12. Kartal SB, Baskurt O, Yavuz AY, Aydin MV. Prognostic value of peripheral blood inflammatory markers in chronic subdural hematoma patients under 65 years. Bratislava Medical Journal 2025;126:1509–17. [CrossRef]
13. Pan W, Zhao D, Zhang C, Li W, Yu J, Wang S, et al. Application of neutrophil/lymphocyte ratio in predicting coronary blood flow and mortality in patients with ST-elevation myocardial infarction undergoing percutaneous coronary intervention. J Cardiol 2015;66:9–14. [CrossRef]
14. Zhao L, Dai Q, Chen X, Li S, Shi R, Yu S, et al. Neutrophil-to-lymphocyte ratio predicts length of stay and acute hospital cost in patients with acute ischemic stroke. J Stroke Cerebrovasc Dis 2016;25:739–44. [CrossRef]
15. Olivo S, Buoite Stella A, Pavan S, Cegalin M, Furlanis G, Cheli M, et al. Admission neutrophil-to-lymphocyte ratio predicts length of hospitalization and need for ICU admission in adults with Status Epilepticus. Seizure 2023;106:80–4. [CrossRef]
16. Kurtulmus A, Yildiz FBP, Boyraz RKS, Celik Z, Gormez A. Inflammatory ratios as predictors of length of hospitalization in psychiatric patients: A multicenter study. Eur Arch Psychiatry Clin Neurosci 2026;276:77–86. [CrossRef]
17. Zhang Y, Yang Y, Long S, Li G. Assessment of peripheral blood cell inflammatory markers in patients with chronic subdural hematoma. Clin Neurol Neurosurg 2020;191:105738. [CrossRef]

DOI: 10.14744/cm.2026.36002
Compreh Med 2026;18(Suppl. 1):119-125

RESEARCH ARTICLE

Patient-Reported Outcomes Following Transsphenoidal Surgery for Acromegaly: A Cross-Sectional Study of Quality of Life, Sleep Quality, and Body Image

Ozan Başkurt[1], Nurperi Gazioğlu[2]

[1]Department of Neurosurgery, Acıbadem Maslak Hospital, İstanbul, Türkiye
[2]İstanbul University-Cerrahpaşa, Pituitary Center, İstanbul, Türkiye

ABSTRACT

Objective: Acromegaly is associated with substantial physical and psychosocial morbidity that may persist despite biochemical disease control. This study aimed to evaluate patient-reported quality of life, sleep quality, and body image in patients with acromegaly following transsphenoidal surgery and to examine their relationships with biochemical remission.

Materials and Methods: This single-center cross-sectional study included 36 adults who underwent transsphenoidal surgery for acromegaly. Quality of life was assessed using the Acromegaly Quality of Life Questionnaire (AcroQoL), with preoperative scores retrospectively reconstructed from patient recall and postoperative scores assessed at the time of the study. Sleep quality and body image were evaluated using the Pittsburgh Sleep Quality Index (PSQI) and Body Image Scale (BIS), respectively. Biochemical remission was defined according to postoperative GH and IGF-1 levels. Nonparametric tests, effect sizes, and Spearman correlation analyses were used.

Results: Postoperative normalized AcroQoL scores were higher than retrospectively recalled preoperative scores (80.11±17.34 vs. 60.64±20.85; $p<0.001$; effect size $r=0.77$). Patients who achieved biochemical remission had higher postoperative AcroQoL scores than those without remission (88.80±10.95 vs. 74.03±18.59; $p=0.013$). In contrast, PSQI and BIS scores did not differ significantly according to remission status. Poorer sleep quality was moderately associated with lower postoperative AcroQoL scores (Spearman $\rho=-0.39$; $p=0.018$). Postoperative follow-up duration was not significantly correlated with AcroQoL, PSQI, or BIS scores.

Conclusion: Higher postoperative patient-reported quality-of-life scores were observed, particularly among patients with biochemical remission. However, sleep disturbances and body image concerns frequently persisted, supporting the routine use of patient-reported outcome measures alongside biochemical assessment during postoperative follow-up.

Keywords: Acromegaly, body image, pituitary adenoma, quality of life, sleep, transsphenoidal surgery

How to cite this article: Başkurt O, Gazioğlu N. Patient-Reported Outcomes Following Transsphenoidal Surgery for Acromegaly: A Cross-Sectional Study of Quality of Life, Sleep Quality, and Body Image. Compreh Med 2026;18(Suppl. 1):119-125.

INTRODUCTION

Acromegaly is a rare chronic endocrine disorder caused by prolonged exposure to excessive growth hormone (GH), most commonly resulting from a GH-secreting pituitary adenoma.[1] Persistent GH hypersecretion and elevated insulin-like growth factor-1 (IGF-1) levels induce progressive structural and functional alterations involving multiple organ systems, including the cardiovascular, metabolic, respiratory, and musculoskeletal systems.[2] Owing to its insidious onset and delayed diagnosis, many patients present with irreversible somatic changes, multiple comorbidities, and substantial impairment in health-related quality of life (HRQoL).[3,4]

Address for Correspondence: Ozan Başkurt, Department of Neurosurgery, Acıbadem Maslak Hospital, İstanbul, Türkiye
E-mail: ozanbskrt@gmail.com **ORCID ID:** 0000-0003-1021-6529

Received date: 19.07.2026
Revised date: 01.09.2026
Accepted date: 17.09.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

The primary objective of acromegaly treatment is to achieve biochemical disease control through transsphenoidal surgery, medical therapy, and, when indicated, radiotherapy.[2,5] Transsphenoidal surgery remains the first-line treatment for most patients, and biochemical remission rates have improved considerably over the past decades. Nevertheless, accumulating evidence indicates that biochemical remission alone does not necessarily translate into complete recovery of patients' perceived health status or overall well-being.[4,6] Accordingly, recent international consensus statements recommend complementing biochemical outcomes with patient-reported outcome measures (PROMs) to provide a more comprehensive assessment of treatment success.[1]

Among PROMs, health-related quality of life is the most extensively investigated outcome in acromegaly and is commonly assessed using the disease-specific Acromegaly Quality of Life Questionnaire (AcroQoL).[7] Although significant improvements in quality of life have been reported following surgical treatment, recovery is often incomplete, particularly in psychosocial domains.[6,8] Sleep disturbances, especially obstructive sleep apnea and impaired sleep quality, frequently persist despite biochemical disease control and continue to adversely affect daily functioning and HRQoL.[9–11] Likewise, persistent craniofacial and acral changes may negatively influence body image and psychosocial adjustment even after successful treatment.[12]

Despite increasing recognition of PROMs in acromegaly, studies simultaneously evaluating quality of life, sleep quality, and body image remain scarce. Furthermore, the relationship between these patient-reported outcomes and biochemical remission has not been sufficiently investigated. Therefore, the present study aimed to evaluate quality of life, sleep quality, and body image in patients who had undergone transsphenoidal surgery for acromegaly and to investigate their association with biochemical remission.

MATERIALS and METHODS

Study Design and Participants

This single-center, cross-sectional survey was conducted at the Department of Neurosurgery, Acıbadem Maslak Hospital following approval from the İstinye University Human Research Ethics Committee (date: 04.04.2024; approval no: 24-65). Adult patients who had previously undergone transsphenoidal surgery for acromegaly and were under routine postoperative follow-up were invited to participate during Acromegaly Awareness Day, held on November 1, 2023.[13] All participants provided written informed consent before enrollment.

A total of 36 consecutive patients were included in the study. The diagnosis of acromegaly was established according to internationally accepted clinical, biochemical, and radiological criteria.[1] Biochemical remission was defined as age- and sex-adjusted normalization of serum IGF-1 concentrations with adequate suppression of GH after oral glucose tolerance testing, in accordance with current consensus recommendations.[14]

Data Collection

All data were collected during a single face-to-face interview conducted by the investigators using a standardized questionnaire. Demographic characteristics, clinical information, surgical history, biochemical remission status, and postoperative follow-up duration were recorded.

Quality-of-life assessments referring to the preoperative period were reconstructed retrospectively based on patients' recall, whereas sleep quality and body image were evaluated only for the current postoperative period. Because the study employed a cross-sectional design, no prospective preoperative assessments were available.

Health-Related Quality of Life

HRQoL was evaluated using the AcroQoL, a disease-specific instrument consisting of 22 items that assess physical functioning and psychological well-being, including appearance and personal relationships.[7]

Raw AcroQoL scores range from 22 to 110. To facilitate interpretation and comparison with previous studies, total scores were normalized to a 0–100 scale according to the developers' recommendations, with higher scores indicating better quality of life. Participants completed the questionnaire twice: once reflecting their current postoperative status and once retrospectively recalling their perceived quality of life before surgery.

Sleep Quality

Sleep quality was assessed using the Pittsburgh Sleep Quality Index (PSQI), which evaluates subjective sleep quality over the previous month across seven domains.[15] Total scores range from 0 to 21, with scores >5 indicating poor sleep quality. The validated Turkish version of the PSQI was used.[16]

Body Image

Body image was evaluated using the Body Image Scale (BIS)[17] developed by Secord and Jourard and validated in Turkish by Hovardaoğlu.[18] The scale consists of 40 items scored on a five-point Likert scale, yielding total scores ranging from 40 to 200, with higher scores indicating greater body satisfaction.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics version 27.0 (IBM Corp., Armonk, NY, USA). The distribution of continuous variables was evaluated using graphical methods (histograms and Q–Q plots) together with the Shapiro–Wilk test.

Because questionnaire scores are commonly reported as mean±standard deviation (SD) in the acromegaly literature, these variables were summarized accordingly despite the use of nonparametric statistical methods to facilitate comparison with previous studies. Categorical variables are presented as frequencies and percentages.

Preoperative and postoperative AcroQoL scores were compared using the Wilcoxon signed-rank test. Comparisons between patients with and without biochemical remission were performed using the Mann–Whitney U test. Associations between postoperative AcroQoL and PSQI scores were assessed using Spearman's rank correlation coefficient. In response to the potential heterogeneity associated with postoperative follow-up duration, additional Spearman's rank correlation analyses were performed to assess the relationships between follow-up duration and postoperative AcroQoL, PSQI, and BIS scores.

For nonparametric analyses, effect size (r) was calculated using the formula $r=Z/\sqrt{N}$ and interpreted according to Cohen's criteria (0.10=small, 0.30=medium, ≥0.50=large effect). All statistical tests were two-tailed, and a p-value <0.05 was considered statistically significant.

RESULTS

Study Population

A total of 36 patients who had undergone transsphenoidal surgery for acromegaly were included in the study. The mean age was 46.9±12.5 years, and 23 (63.8%) were female. The mean postoperative follow-up period was 7.3 years (range, 1–19 years). Biochemical remission was achieved in 16 patients (44.4%). Demographic, socioeconomic, and clinical characteristics are presented in Table 1.

The most common presenting symptoms were enlargement of the hands (61.1%), enlargement of the feet (50.0%), facial coarsening (25.0%), snoring or respiratory difficulty (25.0%), and headache (22.2%). Bitemporal hemianopia was present in four patients (11.1%). Mean serum GH levels decreased from 19.70±28.30 ng/mL before surgery to 2.39±3.66 ng/mL postoperatively, whereas mean IGF-1 levels decreased from 614.2±111.5 ng/mL to 188.9±100.9 ng/mL.

Table 1. Demographic, socioeconomic, and clinical characteristics of the study population

Variable	Value
Age, years	46.9±12.5
Female sex, n (%)	23 (63.8)
Marital status, n (%)	
Married	30 (83.3)
Single	5 (13.9)
Widowed	1 (2.8)
Children, n (%)	
Yes	25 (69.4)
No	11 (30.6)
Education, n (%)	
Primary school or below	9 (25.0)
High school	7 (19.4)
University or higher	20 (55.6)
Employment, n (%)	
Employed	19 (52.8)
Unemployed/Retired	17 (47.2)
Years of employment	3.1±1.5
Follow-up duration, years	7.3 (1–19)
Biochemical remission, n (%)	16 (44.4)

Values are presented as mean±standard deviation (SD), median (range), or number (%), as appropriate. SD: Standard deviation

Among the tumors, 30.6% were microadenomas, 61.1% were macroadenomas, and 8.3% were giant adenomas. Knosp grade 3–4 invasion was observed in 38.9% of patients. Endoscopic transsphenoidal surgery was performed in 77.8% of patients, while 22.2% underwent microscopic transsphenoidal surgery. Sparsely granulated somatotroph adenoma was the most common histopathological subtype (45.8%). Reoperation was required in 16.7% of patients.

Quality of Life

Normalized AcroQoL scores were significantly higher in the postoperative assessment than in the retrospectively reconstructed preoperative assessment (80.11±17.34 vs. 60.64±20.85; Wilcoxon signed-rank test, p<0.001). The magnitude of this difference was large (effect size r=0.77), indicating a substantial difference between retrospectively recalled preoperative and current postoperative patient-reported quality-of-life scores (Table 2, Fig. 1). The median paired improvement in the normalized AcroQoL score was 18.2 points (Hodges–Lehmann estimate; 95% CI: 12.5–24.8).

Table 2. Comparison of normalized AcroQoL scores

Outcome	Mean±SD	p	Effect size (r)
Preoperative AcroQoL	60.64±20.85	<0.001*	0.77
Postoperative AcroQoL	80.11±17.34		
No biochemical remission	74.03±18.59	0.013†	0.43
Biochemical remission	88.80±10.95		

*: Wilcoxon signed-rank test; †: Mann–Whitney U test. Effect size interpreted according to Cohen's criteria. AcroQoL: Acromegaly quality of life questionnaire; SD: Standard deviation

Postoperative AcroQoL scores were significantly higher among patients who achieved biochemical remission than among those without remission (88.80±10.95 vs. 74.03±18.59; Mann–Whitney U test, p=0.013). The corresponding effect size was r=0.43, indicating a moderate-to-large between-group difference.

To further characterize patient-reported quality of life, AcroQoL domain scores were analyzed separately. Significant postoperative improvements were observed in the physical, appearance, and personal relationship domains. The greatest improvement was observed in the physical domain, whereas improvements in appearance-related and interpersonal domains were more modest.

Table 3. Comparison of postoperative sleep quality and body image according to biochemical remission

Variable	No remission	Biochemical remission	p
PSQI	5.80±3.61	5.06±2.82	0.496
BIS	130.30±24.37	133.31±33.74	0.767

Values are presented as mean±standard deviation. PSQI: Pittsburgh sleep quality index; BIS: Body image scale

Sleep Quality

The mean postoperative PSQI score was 5.47±3.26. Although patients with biochemical remission demonstrated lower PSQI scores than patients without remission (5.06±2.82 vs. 5.80±3.61), this difference was not statistically significant (p=0.496). The effect size was small (r=0.12) (Table 3).

Body Image

The mean BIS score was 131.64±28.51. No statistically significant difference was observed between patients with and without biochemical remission (133.31±33.74 vs. 130.30±24.37; p=0.767). The corresponding effect size was r=0.05, indicating a negligible difference (Table 3).

Nineteen patients (52.8%) had body image scores above 135, suggesting generally favorable body image perception. The most frequently reported areas of dissatisfaction

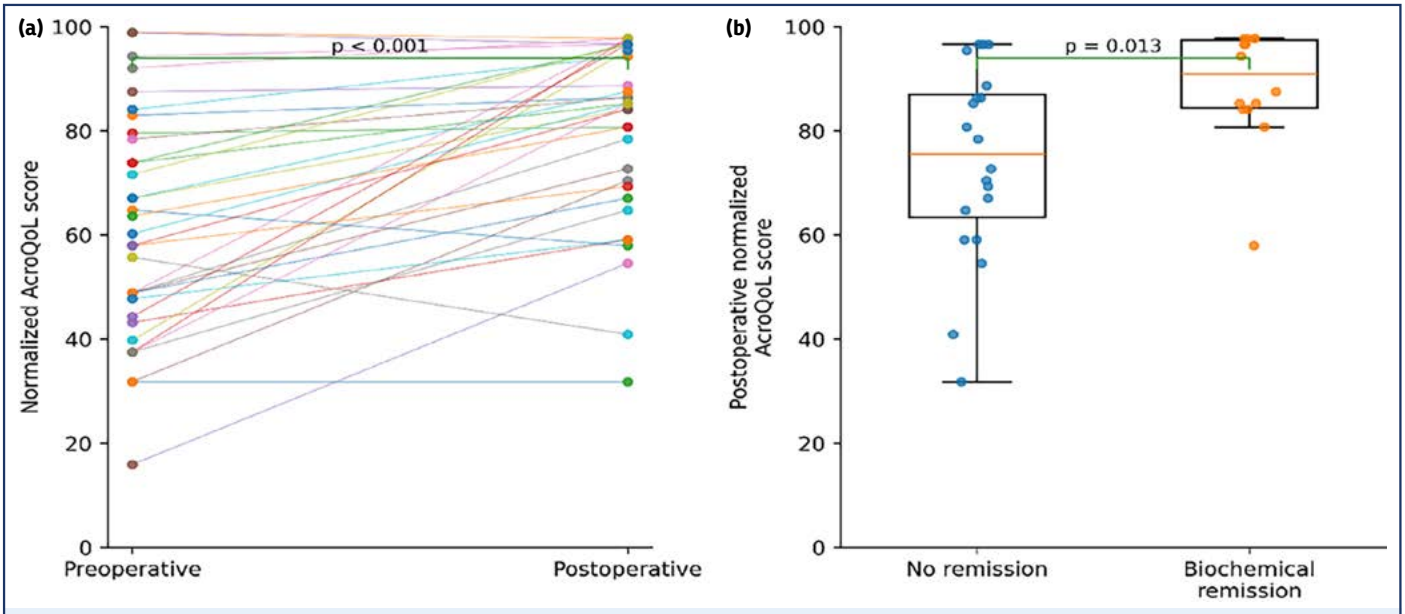


Figure 1. Patient-reported quality-of-life outcomes following transsphenoidal surgery for acromegaly. **(a)** Paired preoperative and postoperative normalized AcroQoL scores. Each line represents an individual patient. Postoperative scores were significantly higher than retrospectively reconstructed preoperative scores (Wilcoxon signed-rank test, p<0.001). **(b)** Distribution of postoperative normalized AcroQoL scores according to biochemical remission status. Patients with biochemical remission had significantly higher scores than those without remission (Mann–Whitney U test, p=0.013)

were sleep pattern, nose, sexual functioning, body posture, and voice.

Correlation Analysis

Spearman correlation analysis demonstrated a moderate negative correlation between postoperative AcroQoL and PSQI scores (ρ=−0.39, p=0.018), indicating that poorer sleep quality was associated with lower quality of life.

Postoperative follow-up duration was not significantly correlated with postoperative AcroQoL (Spearman's ρ=0.12, p=0.483), PSQI (ρ=−0.10, p=0.569), or BIS scores (ρ=0.03, p=0.859).

DISCUSSION

Acromegaly is increasingly recognized as a chronic disease in which treatment success extends beyond biochemical remission.[4,19] Although normalization of GH and IGF-1 concentrations remains the principal therapeutic goal, contemporary consensus statements emphasize that patient-reported outcome measures should complement biochemical assessment.[1,2] This broader perspective is particularly relevant because persistent comorbidities and irreversible somatic changes may continue to affect patients even after successful treatment.[20–22] Accordingly, the present study simultaneously evaluated health-related quality of life, sleep quality, and body image in patients who had undergone transsphenoidal surgery for acromegaly.

The principal finding was the marked difference between retrospectively reconstructed preoperative and current postoperative AcroQoL scores. The approximately 20-point difference in normalized scores was accompanied by a large effect size, indicating a substantial difference in patient-perceived quality of life. Previous studies have similarly reported improved quality of life after transsphenoidal surgery.[8] A systematic review and meta-analysis also demonstrated overall improvements in symptoms and health-related quality of life following treatment of acromegaly.[23] However, quality-of-life recovery is not necessarily complete or uniform. AcroQoL captures physical functioning, appearance, and personal relationships,[7] and previous studies have demonstrated that quality of life may remain impaired despite treatment.[24]

Importantly, the preoperative AcroQoL scores in our study were reconstructed retrospectively from patient recall rather than obtained prospectively. Given the wide range of postoperative follow-up durations, current clinical status and adaptation over time may have influenced patients' recollection of their preoperative condition. Therefore, the observed preoperative–postoperative difference should be interpreted as patient-perceived change rather than as a prospective longitudinal estimate of the effect of surgery. This distinction limits causal interpretation of the observed difference and underscores the need for prospective studies using standardized preoperative and postoperative assessments.

Patients who achieved biochemical remission had significantly higher postoperative AcroQoL scores than those without remission, with a moderate-to-large effect size. This finding is consistent with previous evidence associating long-term biochemical control with better quality of life[6] and disease activity with cognitive and psychological outcomes.[19] Nevertheless, hormonal normalization may not eliminate chronic cardiovascular and metabolic disease, arthropathy, persistent pain, established craniofacial and acral changes, or psychological morbidity.[4,12] Thus, postoperative quality of life likely reflects the combined influence of biochemical activity, chronic comorbidities, psychological status, and social functioning rather than biochemical remission alone.

Sleep quality remained impaired in a substantial proportion of the cohort. The mean postoperative PSQI score was above the commonly accepted threshold for poor sleep quality, and PSQI scores did not differ significantly according to biochemical remission. Sleep-disordered breathing is highly prevalent in acromegaly and may be only partially reversible after biochemical control because of persistent upper-airway changes, craniofacial remodeling, obesity, and associated cardiometabolic disease.[9,25] The moderate negative correlation between PSQI and postoperative AcroQoL further indicated that poorer sleep quality was associated with lower quality of life. Similar associations have been reported in patients with pituitary adenomas.[10] These findings support the systematic assessment of sleep symptoms during follow-up; however, because PSQI was assessed only postoperatively, no conclusion can be drawn regarding the effect of surgery on sleep quality.

Body image was similarly not explained by biochemical remission alone. BIS scores did not differ significantly between remission groups, and although more than half of the patients had scores compatible with generally favorable body image, dissatisfaction persisted in areas including nasal appearance, body posture, sexual functioning, and voice. Previous work suggests that body image perception in acromegaly may be more closely associated with depressive symptoms than with objectively measured acromegalic features.[26] In addition, established skeletal and craniofacial changes may

be only partially reversible. These observations suggest that body image reflects an interaction among persistent physical changes, psychological adaptation, self-esteem, bodily function, and social experience, supporting multidisciplinary assessment when clinically indicated.

The present findings have practical implications for long-term follow-up. Biochemical remission remains an important therapeutic target and was associated with better postoperative quality of life in our cohort. Nevertheless, biochemical control should not be assumed to represent complete patient-perceived recovery. Validated PROMs can identify persistent sleep disturbance, body image concerns, psychological distress, and functional limitations that may not be apparent from endocrine or radiological assessment alone. Incorporating instruments such as AcroQoL and targeted assessments of sleep and body image may therefore provide a more comprehensive evaluation of treatment outcomes.

Several limitations should be considered. The single-center, cross-sectional design and relatively small sample size limit generalizability and preclude causal inference. Most importantly, preoperative AcroQoL scores were retrospectively reconstructed from patient recall, introducing substantial potential for recall bias. The wide range of postoperative follow-up durations may also introduce heterogeneity related to treatment era, additional therapies, comorbidities, and psychosocial adaptation. In additional analyses, however, follow-up duration was not significantly correlated with postoperative AcroQoL, PSQI, or BIS scores. Reliable patient-level data regarding symptom duration before diagnosis, duration of active disease before surgery, and cumulative exposure to GH/IGF-1 excess were not consistently available. This is relevant because prolonged hormonal exposure may contribute to irreversible somatic changes, chronic comorbidities, and persistent impairment in quality of life and body image. Furthermore, PSQI and BIS were assessed only postoperatively, preventing longitudinal evaluation of sleep quality and body image. The absence of healthy or medically treated comparison groups and potential residual confounding by depression, anxiety, obesity, obstructive sleep apnea, pain, and other comorbidities should also be acknowledged.

Despite these limitations, this study simultaneously evaluated three complementary patient-reported outcome domains using validated instruments and focused specifically on surgically treated patients. The inclusion of effect sizes and the additional assessment of follow-up duration strengthen the interpretation of the findings. Prospective multicenter studies incorporating standardized preoperative PROMs, repeated postoperative measurements, objective sleep assessment, and psychological profiling are needed to better define the longitudinal course and determinants of patient-reported outcomes in acromegaly.

CONCLUSION

Transsphenoidal surgery was associated with improved patient-reported quality of life in patients with acromegaly, particularly among those who achieved biochemical remission. Nevertheless, persistent sleep disturbances and body image concerns suggest that successful hormonal control does not necessarily translate into complete recovery of patients' perceived health status. Therefore, postoperative assessment should extend beyond biochemical remission and routinely incorporate validated patient-reported outcome measures to provide a more comprehensive evaluation of treatment success. Future prospective longitudinal studies are needed to confirm these findings and further define the long-term trajectory of patient-reported outcomes after treatment.

Disclosures

Ethics Committee Approval: The study was approved by the İstinye University Human Research Ethics Committee (No: 24-65, Date: 04/04/2024).

Informed Consent: Written informed consent was obtained from all participants before enrollment in the study.

Conflict of Interest Statement: The authors declare that there is no conflict of interest to disclose.

Funding: The authors declare that the study received no funding.

Use of AI for Writing Assistance: OpenAI's ChatGPT was used in July 2026 for English-language editing, grammatical revision, and improvement of the clarity and organization of the manuscript. The authors critically reviewed and revised all AI-assisted content and take full responsibility for the accuracy, integrity, and final content of the manuscript. The AI tool was not used for data collection, statistical analysis, interpretation of the study findings, or generation of scientific conclusions.

Author Contributions: Concept – O.B., N.G.; Design – O.B., N.G.; Supervision – N.G.; Funding – O.B.; Data collection and/or processing – O.B., N.G.; Data analysis and/or interpretation – O.B.; Literature search – O.B., N.G.; Writing – O.B., N.G.; Critical review – O.B., N.G.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Melmed S, di Filippo L, Fleseriu M, Mercado M, Karavitaki N, Gurnell M, et al. Consensus on acromegaly therapeutic outcomes: an update. Nat Rev Endocrinol 2025;21:718–37. [CrossRef]
2. Fleseriu M, Langlois F, Lim DST, Varlamov EV, Melmed S. Acromegaly: pathogenesis, diagnosis, and management. Lancet Diabetes Endocrinol. 2022;10:804–26. [CrossRef]
3. Rowles SV, Prieto L, Badia X, Shalet SM, Webb SM, Trainer PJ. Quality of life (QOL) in patients with acromegaly is severely impaired: use of a novel measure of QOL: acromegaly quality of life questionnaire. J Clin Endocrinol Metab 2005;90:3337–41. [CrossRef]
4. Kocaman BB, Cicek Y, Sahin S, Dogan AB, Uysal S, Muradov I, et al. Psychological profiles of patients with acromegaly: personality traits, quality of life, and clinical correlates. Pituitary 2025;29:24. [CrossRef]
5. Katznelson L, Laws ER Jr, Melmed S, Molitch ME, Murad MH, Utz A, et al. Acromegaly: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 2014;99:3933–51. [CrossRef]
6. Kimball A, Dichtel LE, Yuen KCJ, Woodmansee WW, Haines MS, Nachtigall LB, et al. Quality of life after long-term biochemical control of acromegaly. Pituitary 2022;25:531–9. [CrossRef]
7. Webb SM, Prieto L, Badia X, Albareda M, Catalá M, Gaztambide Set al. Acromegaly Quality of Life Questionnaire (ACROQOL) a new health-related quality of life questionnaire for patients with acromegaly: development and psychometric properties. Clin Endocrinol (Oxf) 2002;57:251–8. [CrossRef]
8. Gu J, Xiang S, He M, Wang M, Gu Y, Li L, et al. Quality of Life in Patients with Acromegaly before and after Transsphenoidal Surgical Resection. Int J Endocrinol 2020;2020:5363849. [CrossRef]
9. Langlois F, Suarez GM, Fleseriu M. Updates in rare and not-so-rare complications of acromegaly: focus on respiratory function and quality of life in acromegaly. F1000Res 2020;9 Faculty Rev-791. [CrossRef]
10. Sagan KP, Andrysiak-Mamos E, Tyburski E, Sagan LM, Syrenicz A. Quality of Life and sleep in patients with pituitary adenoma in relation to tumor type and compression of the optic chiasm. J Clin Med 2021;10:1879. [CrossRef]
11. Kim SH, Kim N, Min KT, Kim EH, Oh H, Choi SH. Sleep disturbance and delirium in patients with acromegaly in the early postoperative period after transsphenoidal pituitary surgery. Medicine (Baltimore) 2020;99(45). [CrossRef]
12. Geraedts VJ, Dimopoulou C, Auer M, Schopohl J, Stalla GK, Sievers C. Health outcomes in acromegaly: depression and anxiety are promising targets for improving reduced quality of life. Front Endocrinol (Lausanne) 2015;5:229. [CrossRef]
13. Worldwide Adrenal and Pituitary Organizations (WAPO). Acromegaly Awareness Day. Available at: https://www.wapo.org/event/acromegaly-awareness-day/ Accessed September 22, 2026.
14. Giustina A, Uygur MM, Frara S, Barkan A, Biermasz NR, Chanson P, et al. Pilot study to define criteria for Pituitary Tumors Centers of Excellence (PTCOE): results of an audit of leading international centers. Pituitary 2023;26:583–96. [CrossRef]
15. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989;28:193–213. [CrossRef]
16. Agargün MY, Kara H, Anlar O. Pittsburgh Uyku Kalitesi Indeksinin Geçerligi ve Güvenirligi. Türk Psikiyatri Dergisi 1996;7:107–15. [Article in Turkish]
17. Secord PF, Jourard SM. The appraisal of body-cathexis: body-cathexis and the self. J Consult Psychol 1953;17:343–7. [CrossRef]
18. Hovardaoğlu S. Vücut algısı ölçeği. 3P Psikiyatri Psikoloji Psikofarmakoloji Dergisi 1993;1(Ek sayı:26–7. [Article in Turkish]
19. Uysal S, Kara Z, Sulu C, Sahin S, Demir AN, Gungormus E, et al. Impact of disease activity on cognitive and psychological outcomes in acromegaly: A prospective study. Growth Horm IGF Res 2025;81:101650. [CrossRef]
20. Dal J, Feldt-Rasmussen U, Andersen M, Kristensen LØ, Laurberg P, Pedersen L, et al. Acromegaly incidence, prevalence, complications and long-term prognosis: a nationwide cohort study. Eur J Endocrinol 2016;175:181–90. [CrossRef]
21. Can M, Kocabaş M, Çordan İ, Çalışkan Burgucu H, Karaköse Met al. Prevalence of comorbidities and associated factors in acromegaly patients in the Turkish population. Turk J Med Sci 2021;51:1146–52. [CrossRef]
22. Demir AN, Sulu C, Kara Z, Sahin S, Ozaydin D, Sonmez O, et al. Changing presentation of acromegaly in half a century: a single-center experience. Pituitary 2023;26:573–82. [CrossRef]
23. Broersen LHA, Zamanipoor Najafabadi AH, Pereira AM, Dekkers OM, van Furth WR, Biermasz NR. Improvement in symptoms and health-related quality of life in acromegaly patients: a systematic review and meta-analysis. J Clin Endocrinol Metab 2021;106:577–87. [CrossRef]
24. Fleseriu M, Fogelfeld L, Gordon MB, Sisco J, Crosby RD, Ludlam WH, et al. An evaluation of the Acromegaly Treatment Satisfaction Questionnaire (Acro-TSQ) in adult patients with acromegaly, including correlations with other patient-reported outcome measures: data from two large multicenter international studies. Pituitary 2020;23:347–58. [CrossRef]
25. Davi' MV, Dalle Carbonare L, Giustina A, Ferrari M, Frigo A, Lo Cascio V, et al. Sleep apnoea syndrome is highly prevalent in acromegaly and only partially reversible after biochemical control of the disease. Eur J Endocrinol 2008;159:533–40. [CrossRef]
26. Dimopoulou C, Leistner SM, Ising M, Schneider HJ, Schopohl J, Rutz S, et al. Body image perception in acromegaly is not associated with objective acromegalic changes but depends on depressive symptoms. Neuroendocrinology 2017;105:115–22. [CrossRef]

DOI: 10.14744/cm.2026.70894
Compreh Med 2026;18(Suppl. 1):126-130

CASE REPORT

Atypical Presentation: A Clear Cell Carcinoma Case Diagnosed with Bone Marrow Metastasis

Dilek Keskin[1], Ahmet Can Toksöz[2], Taşkın Erkinüresin[3], Murat Akarsu[2]

[1]Department of Hematology, University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
[2]Department of Internal Medicine, University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
[3]Department of Pathology, University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye

ABSTRACT

Clear cell renal cell carcinoma (ccRCC) is the most common malignant neoplasm of the kidney. This tumor most frequently metastasizes to the lungs, bones, lymph nodes, and liver, respectively. However, bone marrow metastasis of ccRCC is extremely rare, with only a few cases reported in the literature. In this study, we present a case of ccRCC that initially manifested with bone marrow metastasis, leading to the diagnosis.

Keywords: Biopsy, bone marrow, carcinoma, hemophagocytic, leukemoid reaction, lymphohistiocytosis, renal cell

How to cite this article: Keskin D, Toksöz AC, Erkinüresin T, Akarsu M. Atypical Presentation: A Clear Cell Carcinoma Case Diagnosed with Bone Marrow Metastasis. Compreh Med 2026;18(Suppl. 1):126-130.

INTRODUCTION

A 66-year-old male patient with no known chronic medical history presented to the emergency department with complaints of fever, back pain, fatigue, dyspnea, and nausea. On physical examination, he appeared cachectic and in poor general condition. His temperature was 37.8°C, and he exhibited tachycardia. A firm, slightly painful mass approximately 10 cm in size was palpated in the left lower abdomen. Decreased breath sounds were noted in the left lower hemithorax, with bilateral rales observed in the mid-zones. Informed consent was obtained.

Laboratory findings revealed neutrophilic leukocytosis (18.8×10^3/µL), elevated serum lactate dehydrogenase (LDH), and increased acute-phase reactants. Hemoglobin (Hb) was 14.3 g/dL, LDH was 975 U/L, C-reactive protein (CRP) was 239 mg/L, and INR was 1.26. Renal and hepatic function tests were within normal limits (Table 1 presents the laboratory characteristics).

Contrast-enhanced thoracic CT showed nodular lesions in the lungs, pleural effusion, and bilateral mosaic attenuation. Abdominal CT revealed a heterogeneous mass in the left kidney upper pole, approximately 12 cm in size, with central necrotic areas (Fig. 1).

The patient was hospitalized with presumptive diagnoses of pneumonia and malignancy. Although antibiotic therapy was initiated, fever and inflammatory markers continued to rise. A peripheral blood smear revealed a leukocytosis with a leukoerythroblastic picture. While ferritin levels showed a progressive increase during follow-up (rising from an initial value of 2.085 ng/mL to 8.832 ng/mL), fibrinogen levels demonstrated a downward trend, decreasing from 774 mg/dL to 398 mg/dL, in contrast to the increasing CRP levels. Diagnostic and therapeutic thoracentesis was performed, showing a transudative pleural effusion, and cytological examination revealed mesothelial cells and reactive lymphocytes.

As the patient's general condition worsened, and due to the presence of widespread metastatic disease, further diagnostic procedures could not be performed. Based on the available clinical and laboratory data, bone marrow aspiration and trephine biopsy were conducted with a pre-diagnosis of

Address for Correspondence: Dilek Keskin, Department of Hematology, University of Health Sciences, Kanuni Sultan Süleyman Training and Research Hospital, İstanbul, Türkiye
E-mail: dilekkeskin84@hotmail.com **ORCID ID:** 0000-0002-5332-3538

Received date: 19.09.2025
Revised date: 23.01.2026
Accepted date: 29.01.2026
Online date: 14.10.2026

Comprehensive Medicine published by Kare Media.
OPEN ACCESS This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

Table 1. Laboratory findings of the patient with references

Parameter*	Result	Reference range**
WBC (mm^3)	24.950	4.000–10.000 /mm^3
Neu (mm^3)	23.490	2.000–7.500 /mm^3
Hb (g/dl)/Hct (%)	11.8/38.9	Male: 13.5–17.5 g/dL/41–53%, Female: 12–16 g/dL/36–46%
MCV (fL)	86.4	80–100 fL
Plt (mm^3)	366.000	150.000–450.000 /mm^3
CRP (mg/dl)	257.02	<0.5 mg/dL
Ferritin (ng/mL)	8.532	Male: 30–400 ng/mL, Female: 15–150 ng/mL
Folate/Vit B12	2.9/131	Folate: 3–17 ng/mL, vit B12: 200–900 pg/mL
Na/K (mmol/L)	132/5.2	Na: 135–145 mmol/L, K: 3.5–5.1 mmol/L
BUN/Creatinine	26.52/0.82	BUN: 7–20 mg/dL, Creatinine: 0.6–1.3 mg/dL
AST/ALT (U/L)	265/94	AST: 5–40 U/L, ALT: 5–45 U/L
LDH (U/L)	2055	125–220 U/L
GGT/ALP (U/L)	568/684	GGT: 9–48 U/L, ALP: 44–147 U/L
Bilirubin (mg/dL)	0.81	0.2–1.2 mg/dL
PT/APTT (sec)	18.6/33.6	PT: 11–15 sec, APTT: 25–35 sec
Fibrinogen/D-dimer	398/4.4	Fibrinogen: 200–400 mg/dL, D-dimer: <0.5 µg/mL
ESR (mm/h)	121	Male: <15 mm/h, Female: <20 mm/h
Total Protein/Albumin (g/dL)	5.2/2.3	Total protein: 6.0–8.3 g/dL, Albumin: 3.5–5.0 g/dL

WBC: White blood cell; Neu: Neutrophil; Hb: Hemoglobin; HCT: Hematocrit; MCV: Mean corpuscular volume; Plt: Platelet; CRP: C-reactive protein; BUN: Blood urea nitrogen; AST: Aspartate aminotransferase; ALT: Alanine aminotransferase; LDH: Lactate dehydrogenase; GGT: Gamma-glutamyl transferase; ALP: Alkaline phosphatase; PT: Prothrombin time; APTT: Activated partial thromboplastin time; ESR: Erythrocyte sedimentation rate. Reference ranges are adapted from standard clinical biochemistry sources (Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, 6th edition; Henry's Clinical Diagnosis and Management by Laboratory Methods, 23rd edition).

hemophagocytosis. The aspirate yielded a dry tap, and imprint smears showed rosette-like clusters of cells and an increased number of histiocytes (Fig. 2). Microscopic examination of hematoxylin and eosin-stained sections revealed a nodular infiltration composed of large, round cells with clear cytoplasm, prominent nucleoli, and centrally located nuclei. Subsequent immunohistochemical analysis showed that the infiltrative area was negative for MPO, CD61, and Glycophorin A, while it was positive for PanCK, PAX2, and Vimentin. The lesion was negative for CD10, Calretinin, RCC, EMA, WT-1, CK5/6, E-Cadherin, Cyclin D1, Bcl-2, CD3, CD20, CD30, and ALK1. In light of these findings, the case was diagnosed as a carcinoma metastasis. Considering the patient's clinical history of a renal mass, the morphological features of large round cells with clear cytoplasm, prominent nucleoli, and centrally located nuclei on hematoxylin and eosin staining, along with the immunohistochemical positivity for PanCK, PAX2, and Vimentin, the findings were interpreted as being most consistent with metastatic clear cell RCC (Fig. 3).

Empirical dexamethasone (4 mg/day) was initiated due to an HLH-like clinical presentation, along with supportive hemo-

Figure 1. The abdominal CT findings of the case are shown below. Abdominal CT demonstrating a 12 cm heterogeneous contrast-enhancing mass in the left renal fossa

CT: Computed tomography

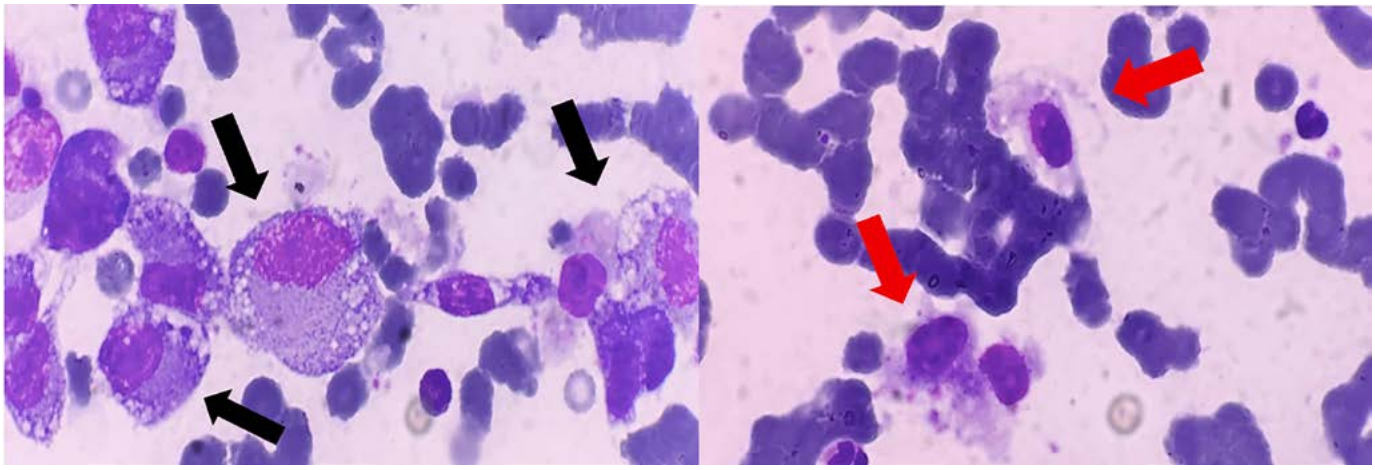


Figure 2. The bone marrow aspiration smear of the patient is depicted. Bone marrow aspiration smear showing large non-hematologic cells with clear cytoplasmic vacuoles (black arrows) and increased histiocytic activity (red arrows) (H&E, ×100)

static therapy. However, due to the development of respiratory failure, oncological treatment could not be administered, and intensive care support was started. Unfortunately, the patient passed away due to bacterial septicemia.

DISCUSSION

Solid organ neoplasms with bone marrow metastasis can dominate the clinical and laboratory findings through well-known mechanisms. First, if untreated, bone marrow metastasis may trigger hemophagocytic syndrome. Second, medullary narrowing and secondary reticulin fibrosis may also lead to myelofibrotic anemia.[1] The occurrence of bone marrow metastasis in clear cell renal cell carcinoma is extremely rare but can mimic hematological malignancies in laboratory findings.[2] In this case, the bone marrow infiltration presented with leukocytosis, anemia, elevated serum LDH, and fever, along with progressively increasing ferritin levels and a gradual decline in fibrinogen levels, contributing to the clinical picture. Therefore, considering the possibility of an HLH-like presentation in the differential diagnosis, and given that this presentation constitutes a life-threatening hematologic–immunologic emergency, early recognition and timely intervention were considered essential, and a bone marrow aspiration and biopsy were performed.

Renal cell carcinoma itself can cause fever, and in cases where pathological sampling from the primary tumor or distant metastases is not feasible, bone marrow sampling can provide a diagnostic solution.[3] Furthermore, RCC can induce a leukemoid reaction due to paraneoplastic granulocyte-colony stimulating factor (G-CSF) secretion, which, along with the hyperinflammatory state created by the disease, can lead to leukemoid reactions.[4]

In terms of oncopathological staging, solid organ neoplasms with bone marrow metastasis are classified as stage IV in the TNM staging system.[5] When advanced-stage cases reach a medically inoperable point, clinicians may alter their approach. This often results in the avoidance of invasive diagnostic procedures. Bone marrow aspiration and trephine biopsy can be valuable in such cases, as they are relatively simple bedside procedures that can be performed by appropriately trained clinicians involved in the patient's care.

The prognostic significance of bone marrow involvement has been well documented in the literature, as bone marrow micrometastases detected at diagnosis have been associated with poorer outcomes in several solid malignancies, particularly breast cancer, underscoring the importance of bone marrow evaluation from an oncological perspective.[6,7]

Reports of RCC with bone marrow metastasis are extremely rare in the literature. Whether this rarity is due to the infrequency of bone marrow metastasis itself or the academic community's tendency to undervalue case reports remains unclear. Furthermore, the assumption that there is no further action to be taken, coupled with the view that bone marrow aspiration and biopsy would not alter outcomes, likely contributes to their infrequent reporting.[8]

It is well known that most solid tumors can infiltrate bone marrow through micrometastasis, and this may lead to cytopenias, inflammatory processes, and hemophagocytosis.[9] Hematology specialists routinely consider bone marrow

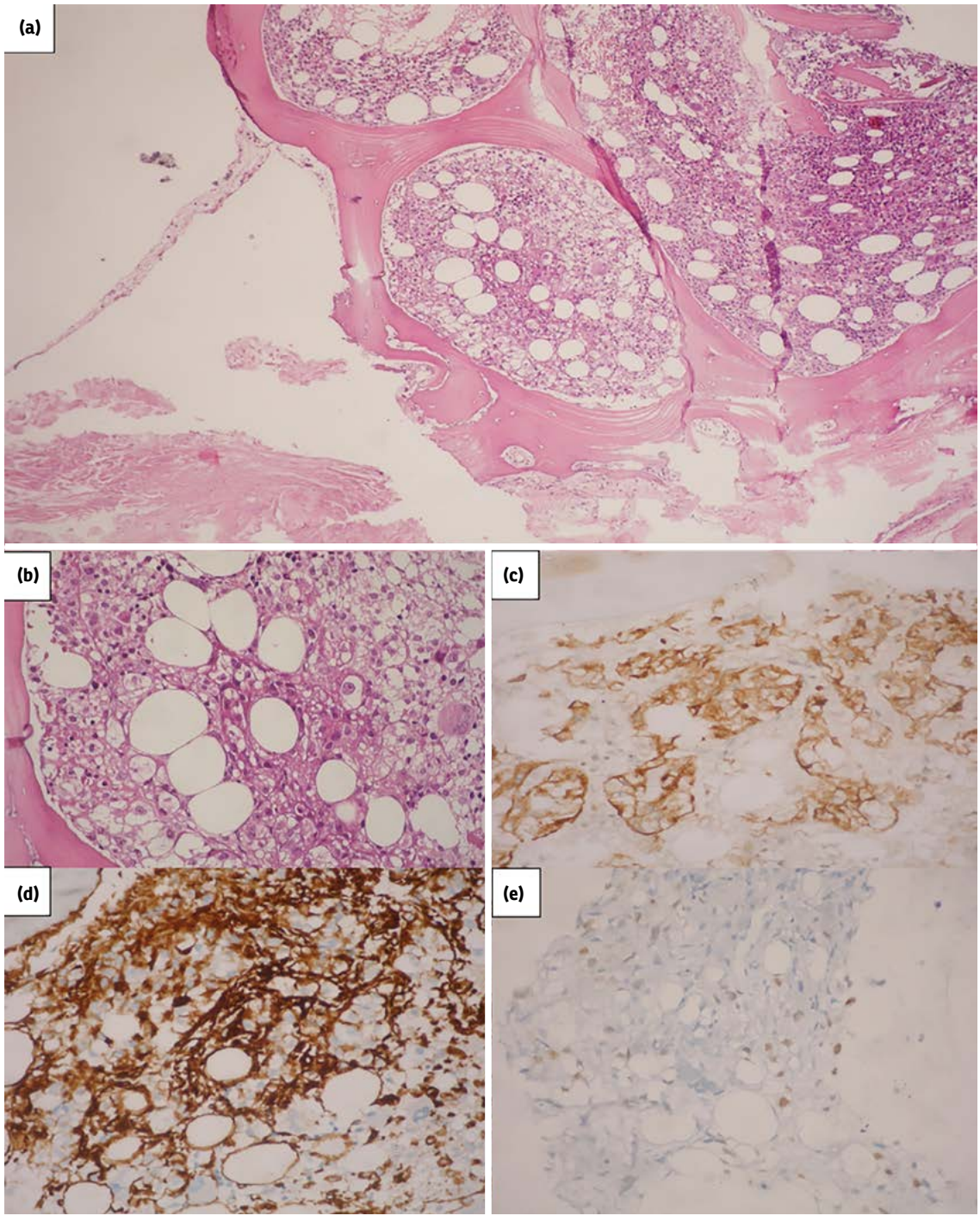


Figure 3. The immunohistochemical findings of the bone marrow trephine biopsy are shown. **(a)** Bone marrow trephine biopsy showing a paratrabecular infiltrate of large, round cells with clear cytoplasm and centrally located nuclei (H&E, ×10). **(b)** High-power view showing an infiltrate of large, round cells with clear cytoplasm, centrally located nuclei, and prominent nucleoli (H&E, ×40). **(c)** Immunohistochemical staining shows membranous positivity for PanCK in the large, round cells with clear cytoplasm and centrally located nuclei forming the nodule, whereas surrounding bone marrow cells remain negative (PanCK, ×40). **(d)** High-power immunohistochemical view showing suspicious membranous staining for vimentin in the large, round cells with clear cytoplasm and centrally located nuclei forming the nodule (Vimentin, ×40). E. High-power immunohistochemical view demonstrating nuclear positivity for PAX2 in the majority of cells forming the nodule (PAX2, ×40)

trephine biopsy in lymphoma cases for accurate clinical staging, especially in scenarios where disease refractoriness, transplant eligibility, or discordance with PET-CT findings may influence clinical management.[10] Therefore, for confirmed cases of malignancy, particularly when PET-CT fails to detect bone marrow infiltration, bone marrow aspiration and trephine biopsy are critical for accurate staging and management.

CONCLUSION

Bone marrow metastasis is a rare occurrence in renal cell carcinoma cases. In advanced-stage cases where the patient's condition deteriorates, and diagnostic procedures cannot be performed, bone marrow sampling should not be avoided, as it may provide a diagnostic solution. Bone marrow aspiration and biopsy remain valuable tools in the diagnostic workup of RCC with bone marrow involvement.

Disclosures

Ethics Committee Approval: This is a single case report, and therefore ethics committee approval was not required in accordance with institutional policies.

Informed Consent: Written informed consent was obtained from the patient for publication of clinical information and images.

Conflict of Interest Statement: The authors declared no conflicts of interest.

Funding: The authors declared no financial support for this work.

Use of AI for Writing Assistance: The authors declared that they used AI-assisted language editing (ChatGPT) only for linguistic refinement during manuscript preparation. All content has been reviewed and verified by the authors for accuracy and originality.

Author Contributions: Concept – D.K.; Design – D.K., A.C.T.; Supervision – T.E., M.A.; Materials – M.A., T.E.; Data collection and/or processing – D.K., A.C.T.; Data analysis and/or interpretation – D.K.; Literature search – D.K., T.E.; Writing – D.K.; Critical review – D.K.

Peer-review: Externally peer-reviewed.

REFERENCES

1. Rani HS, Hui M, Manasa PL, Uppin SG, Uppin MS, Paul TR, et al. Bone Marrow Metastasis of solid tumors: a study of 174 cases over 2 decades from a single ınstitution in India. Indian J Hematol Blood Transfus 2022;38:8–14. [CrossRef]
2. Chao CT, Kao CC, Lee SY, Ho SJ, Jhuang YJ, Li LH, et al. Renal cell carcinoma with secondary hemophagocytic syndrome: A case report. Can Urol Assoc J 2012;6:E64–6. [CrossRef]
3. Palapattu GS, Kristo B, Rajfer J. Paraneoplastic syndromes in urologic malignancy: the many faces of renal cell carcinoma. Rev Urol 2002;4:163–70.
4. Zhang X, Yan L, Yuan X, Bai T, Zhang L, Han S. Rapid exacerbation featuring acute leukemoid reaction after retrolaparoscopic nephrectomy: a rare case report of renal cell carcinoma with postoperative comprehensive genomic profiling. World J Surg Oncol 2020;18:155. [CrossRef]
5. Kumar S, Kumar K, Agrawal R, Choudhary V. Bone marrow metastasis in non-hematological malignancies: data from a tertiary care hospital in North India. Cureus 2025;17:e82757. [CrossRef]
6. Vincent-Salomon A, Bidard FC, Pierga JY. Bone marrow micrometastasis in breast cancer: review of detection methods, prognostic impact and biological issues. J Clin Pathol 2008;61:570–6. [CrossRef]
7. Janni W, Gastroph S, Hepp F, Kentenich C, Rjosk D, Schindlbeck C, et al. Prognostic significance of an increased number of micrometastatic tumor cells in the bone marrow of patients with first recurrence of breast carcinoma. Cancer 2000;88:2252–9. [CrossRef]
8. Khan S, Awan SA, Jahangir S, Kamran S, Ahmad IN. Bone marrow metastasis in clear cell renal cell carcinoma: a case study. Cureus 2019;11:e4181. [CrossRef]
9. Löfstedt A, Jädersten M, Meeths M, Henter JI. Malignancy-associated hemophagocytic lymphohistiocytosis in Sweden: incidence, clinical characteristics, and survival. Blood 2024;143:233–42. [CrossRef]
10. Guler B, Erdogan EB, Adilli A, Uysal O. Can PET-CT replace bone marrow biopsy for lymphoma staging? Retrospective analysis of 198 Hodgkin and non-Hodgkin lymphoma cases. Ann Med Res 2021;28:1206–12. [CrossRef]