

e-ISSN: 2636-8579

JHSM

Journal of Health Sciences and Medicine

VOLUME: 8

ISSUE: 6

YEAR: 2025

HONORARY EDITOR

Prof. Osman GÜLER

Department of General Surgery, Faculty of Medicine, Kastamonu University,
Kastamonu, Türkiye

EDITORS-IN-CHIEF

Prof. Aydın ÇİFCİ

Department of Internal Medicine, Faculty of Medicine,
Kırıkkale University, Kırıkkale, Türkiye

Prof. Mehmet Özgür ERDOĞAN

Department of Emergency Medicine, Bakırköy Sadi Konuk Training and Research Hospital,
University of Health Sciences, İstanbul, Türkiye

ASSOCIATE EDITORS-IN-CHIEF

Assoc. Prof. Elif PINAR BAKIR

Department of Restorative Dentistry, Faculty of Dentistry,
Dicle University, Diyarbakır, Türkiye

Asst. Prof. Furkan SOY

Department of Orthopedics and Traumatology,
Faculty of Medicine, Kırıkkale University, Kırıkkale, Türkiye

Asst. Prof. Emine SARMAN

Department of Histology and Embriology, Faculty of Medicine,
Afyonkarahisar Health Sciences University, Afyonkarahisar, Türkiye

Assoc. Prof. Tuğba GÜRBÜZ

Department of Obstetrics and Gynecology, Medistate Hospital,
İstanbul, Türkiye

ENGLISH LANGUAGE EDITOR

Assoc. Prof. Mustafa CİVELEKLER

Department of Ophthalmology, Gülhane Training and Research
Hospital, University of Health Sciences, Ankara, Türkiye

STATISTICS EDITOR

Ahsen CEYLAN, MD

Medical Devices Technical Regulation Expert, Clinical Expert, UDEM,
Ankara, Türkiye

EDITORIAL BOARD

Assoc. Prof. Ahmet EKEN

Department of Medical Biology, Faculty of Medicine,
Erciyes University, Kayseri, Türkiye

Prof. Alpaslan TANOĞLU

Department of Gastroenterology, Medical Park Göztepe Hospital
Complex, Faculty of Medicine, Bahçeşehir University,
İstanbul, Türkiye

Prof. Alpaslan Kemal TUZCU

Division of Endocrinology and Metabolism, Department of
Internal Medicine, Faculty of Medicine, Dicle University,
Diyarbakır, Türkiye

Prof. Ayşegül ALTUNKESER

Department of Radiology, Konya City Hospital,
Hamidiye Faculty of Medicine, University of
Health Sciences, Konya, Türkiye

Prof. Berna AKINCI ÖZYÜREK

Department of Chest Diseases, Ankara Atatürk
Sanatorium Training and Research Hospital,
University of Health Sciences,
Ankara, Türkiye

Spec. Bulut DEMİREL, MD

Department of Emergency Medicine, Faculty of Medicine,
Queen Elizabeth University, Paisley,
Glasgow, United Kingdom

Prof. Can CEDİDİ

Department of Plastic, Reconstructif and Aesthetic Surgery,
Bremen, Germany

Prof. Ebru OLGUN

Department of Periodontology, Faculty of Dentistry, Kırıkkale
University, Kırıkkale, Türkiye

Prof. Ekrem ÜNAL

Department of Pediatric Hematology & Oncology,
Medical Point Gaziantep Hospital,
Gaziantep, Türkiye

Prof. Ela CÖMERT

Department of Ear Nose Throat, Faculty of Medicine, Kırıkkale
University, Kırıkkale, Türkiye

Assoc. Prof Emrah ÖZAKAR

Department of Pharmaceutical Technology, Faculty of Pharmacy,
Atatürk University, Erzurum, Türkiye

Assoc. Prof. Faruk PEHLİVANLI

Department of General Surgery, Faculty of Medicine, Kırıkkale
University, Kırıkkale, Türkiye

Prof. Fatma NİŞANCI KILIÇ

Department of Nutrition and Dietetics, Faculty of Health Sciences,
Kırıkkale University, Kırıkkale, Türkiye

Asst. Prof. Fuat TÜRK

Department of Computer Engineering, Faculty of Engineering,
Kırıkkale University, Kırıkkale, Türkiye

Assoc. Prof. Harun DÜĞEROĞLU

Department of Internal Medicine, Faculty of Medicine,
Ordu University, Ordu, Türkiye

Spec. Hidayet MAMMADZADA, MD

Department of Endocrinology and Metabolism, Bakü Medical Plaza
Hospital, Bakü, Azerbaijan

Prof. İbrahim Celalettin HAZNEDAROĞLU

Division of Hematology, Department of Internal Medicine,
Faculty of Medicine, Hacettepe University,
Ankara, Türkiye

Assoc. Prof. İbrahim Halil İNANÇ

Department of Cardiology, Phoenixville Hospital, Towerhealth,
Phoenixville, Pennsylvania, USA

Assoc. Prof. İhsan SOLMAZ

Department of Internal Medicine, Gazi Yaşargil Training
and Research Hospital, University of Health Sciences,
Diyarbakır, Türkiye

Assoc. Prof. İzzet BİNGÖL

Department of Orthopedics and Traumatology, Ankara Bilkent
City Hospital, Ankara, Türkiye

Prof. Kaan OKYAY

Department of Cardiology, Başkent University Ankara Hospital,
Faculty of Medicine, Başkent University, Ankara, Türkiye

Assoc. Prof. Kenan ÇADIRCI

Department of Internal Medicine, Erzurum Region Training
and Research Hospital, Erzurum Faculty of Medicine,
University of Health Sciences, Erzurum, Türkiye

Prof. Mehmet ÇITIRIK

Department of Ophthalmology, Ankara Etlik City Hospital,
Ankara, Türkiye

Assoc. Prof. Mehmet ZENGİN

Department of Medical Pathology, Ankara Training and
Research Hospital, University of Health Sciences,
Ankara, Türkiye

Assoc. Prof. Mehmet Emin DEMİR

Department of Nephrology, Medicana International Ankara
Hospital, Faculty of Medicine, Atılım University, Ankara, Türkiye

Assoc. Prof. Mehmet Fatih YETKİN

Department of Neurology, Faculty of Medicine,
Erciyes University, Kayseri, Türkiye

Prof. Mehmet Sinan DAL

Department of Hematology and Bone Transplantation Unit, Dr.
Abdurrahman Yurtaslan Ankara Oncology Training and Research
Hospital, University of Health Sciences, Ankara, Türkiye

Assoc. Prof. Meltem HENDEK

Department of Periodontology, Faculty of Dentistry,
Kırıkkale University, Kırıkkale, Türkiye

Assoc. Prof. Muhammed KARADENİZ

Department of Cardiology, Faculty of Medicine, Kırıkkale
University, Kırıkkale, Türkiye

Assoc. Prof. Murat DOĞAN

Department of Internal Medicine, Hitit University Erol Olçok
Training and Research Hospital, Faculty of Medicine,
Hitit University, Çorum, Türkiye

Prof. Murat KEKİLLİ

Division of Gastroenterology, Department of Internal Medicine,
Faculty of Medicine, Gazi University, Ankara, Türkiye

Assoc. Prof. Mustafa ÇAPRAZ

Department of Internal Medicine, Faculty of Medicine,
Amasya University, Amasya, Türkiye

Prof. Neven SKITARELIC

Department of Ear Nose Throat, School of Medicine,
University of Zadar, Zadar, Croatia

Prof. Nuray BAYAR MULUK

Department of Ear Nose Throat, Faculty of Medicine,
Kırıkkale University, Kırıkkale, Türkiye

Assoc. Prof. Özge VERGİLİ

Department of Physiotherapy and Rehabilitation, Faculty of Health
Sciences, Kırıkkale University, Kırıkkale, Türkiye

Prof. Ranko MLADINA

Department of Ear Nose Throat, Zagreb, Croatia

Assoc. Prof. Rüçhan Bahadır CELEP

Division of Gastroenterologic Surgery, Department of General
Surgery, Hygeia Hospital, Tirana, Albania

Prof. Salih CESUR

Department of Infectious Diseases and Clinical Microbiology,
Ankara Training and Research Hospital, University of
Health Sciences, Ankara, Türkiye

Assoc. Prof. Süleyman GÖKMEN

Department of Food Processing, Technical Sciences Vocational
School, Faculty of Engineering, Karamanoğlu Mehmetbey
University, Karaman, Türkiye

Assoc. Prof. Tuğçe ŞAHİN ÖZDEMİREL

Department of Chest Diseases, Ankara Atatürk Sanatorium Training
and Research Hospital, University of Health Sciences,
Ankara, Türkiye

Prof. Vedat TOPSAKAL

Department of Ear Nose Throat, Faculty of Medicine,
University Hospital Brussels, Brussels, Belgium

Assoc. Prof. Yaşar TOPAL

Department of Pediatrics, Faculty of Medicine, Muğla Sıtkı
Koçman University, Muğla, Türkiye

Assoc. Prof. Yücel YILMAZ

Department of Cardiology, Kayseri City Hospital,
Kayseri, Türkiye

Assoc. Prof. Zafer PEKKOLAY

Division of Endocrinology and Metabolism,
Department of Internal Medicine, Faculty of Medicine,
Dicle University, Diyarbakır, Türkiye

Prof. Zaim JATIC

Department of Family Medicine, Faculty of Medicine,
University of Sarajevo, Sarajevo,
Bosnia-Herzegovina

Assoc. Prof. Ziya ŞENCAN

Department of Ear Nose Throat, Faculty of Medicine, Kırıkkale
University, Kırıkkale, Türkiye

FRANCHISE OWNER

MediHealth Academy Publishing

LAYOUT EDITOR

Şerife KUTLU

Engineer, MediHealth Academy Publishing, Ankara, Türkiye

CORRESPONDENCE ADDRESS

Sağlık Bilimleri Üniversitesi Teknoloji Geliştirme Merkezi (SBÜ TEKMER)
Gülhane Yerleşkesi, Emrah Mah. General Dr. Tevfik Sağlam Cad. No: 2/67
İç Kapı No: 105, Keçiören/Ankara, Türkiye
E-mail: mha@medihealthacademy.com
www.medihealthacademy.com

ARTICLE SUBMISSION ADDRESS

<https://dergipark.org.tr/tr/journal/2316/submission/step/manuscript/new>

Original Articles

- Is joint involvement in systemic lupus erythematosus always inflammatory? 953-958**
Özsoy Z, Bulat B, Sandal Uzun G, et al.
- Investigation of systemic treatment responses and prognostic factors in metastatic gastroenteropancreatic neuroendocrine carcinomas (GEP-NECs) 959-965**
Türkel A, Ercan Uzundal D, Gürler F, Deliktaş Onur İ, Gülcü S, Yalçıntaş Arslan Ü.
- The impact of perceived social support, internalized stigma, and childhood trauma on quality of life in schizophrenia spectrum disorders: a mediation and moderation approach..... 966-974**
Korkmaz U, Helvacı Çelik FG.
- Global research trends and collaborative networks in renal colic: a bibliometric analysis of the 100 most cited articles 975-980**
Öktem Ç, Siber V.
- Aberrant left hepatic artery arising from the left gastric artery: perioperative effects of a risky vascular anomaly in gastrectomy 981-987**
Kılavuz H, Düzköylü Y, Şevik HO, Tekin O, Yardımcı AH, Yüksel S.
- Comparison of color stability of different 3D-printed resins..... 988-994**
Koca B, Akarsu S, Aktuğ Karademir S, Atasoy S.
- The impact of the COVID-19 pandemic on gastric cancer stage and nutritional status 995-999**
Gül MC, Hafızoğlu E.
- The mediating role of perceived stress: an examination of the relationship between health literacy and healthy living awareness..... 1000-1005**
Yeşildal M, Aydoğdu A, Çelen R, Yaver T.
- Evaluation of carotid artery calcification on panoramic radiographies..... 1006-1010**
Altındağ A, Altındağ Ö.
- Evaluation of sex estimation through mandibular morphometric measurements on cone-beam computed tomography images 1011-1017**
Çelebi E, Ersoy Köleğ SN, Koşkan Ö, Görmez Ö.
- Bicycle accidents and injury patterns in emergency department admissions: a single-center experience..... 1018-1022**
Durdu T, Yıldırım Çelik Y, Ertürk N, Özensoy HS, Altaş F, Oğuztürk H.
- Impact of atrial fibrillation on cognitive decline and cortical atrophy in Alzheimer's disease 1023-1029**
Totuk Ö, Güdek HC, Şahin Ş.
- Evaluation of clinical and laboratory findings in cases of infantile hypertrophic pyloric stenosis 1030-1033**
Kurt F, Kaya M.

CONTENTS

Original Articles

- Mothers' knowledge and attitudes regarding circumcision and the influence of parental demographics: a retrospective study..... 1034-1040**
Erden D, Arıca V.
- Evaluation of the MAPH score in predicting acute ischemic stroke severity and major vessel occlusion in the emergency department..... 1041-1048**
Sarcan E, Yazla M.
- Diagnosis, timing of procedure, and perinatal outcomes in fetocide cases due to fetal anomalies: a five-year experience from Türkiye..... 1049-1054**
Yılmaz EBS, Sağlam E.
- Novel GLA mutation in a high-consanguinity region: insights from Fabry disease screening in Türkiye 1055-1059**
Ada S, Özcan ŞG, Altunören O, Eren N, Toptaş T.
- The effectiveness of low-dose intravenous ketamine on postoperative pain management in arthroscopic knee surgery under spinal anesthesia..... 1060-1064**
Özkavak D, Yılmaz A, Kılıç Y, Özkavak S, Göğüş N.
- Examining jigsaw, metaverse, and classic learning methods' impact on nursing students' self-efficacy in pediatric drug administration: a mixed-methods study..... 1065-1075**
Tutar Ş, Özgörü H, Durna F, Şahin Y.
- The efficacy of transperineal and transabdominal ultrasound findings in predicting spontaneous vaginal delivery during active labor..... 1076-1081**
Büyüknisan E, Yücel N, Yücel O, Aydın ŞG, Yazkan E.
- Bilingual evaluation of YouTube™ videos on teething symptoms and remedies..... 1082-1088**
Demir D, Çerçi Akçay H.
- Separate and combined effects of maternal obesity and preeclampsia on perinatal outcomes: a retrospective four-group analysis..... 1089-1098**
Sağlam E.
- Epidemiology of viral respiratory tract infections in pediatric ICU: a single-center study in Türkiye 1099-1104**
Gülyüz A, Poyraz N, Yıldız Ataş S.
- How reliable are YouTube videos on dental implants? A comparison between Turkish and English content 1105-1111**
Yüceer Çetiner E, Çetiner RB.
- Immunonutritional profiling for prognostic assessment in diffuse large b-cell lymphoma: an initial evaluation of the C-reactive Protein-Albumin-Lymphocyte (CALLY) Index 1112-1117**
Pınar İE, Özkocaman V, Ersal T, et al.

CONTENTS

Original Articles

- Comparison of the perinatal outcomes and antibiotic susceptibility patterns of acute cystitis associated with *Escherichia coli* and *Klebsiella pneumoniae* infections during the second trimester of pregnancy 1118-1123**
Yıldırım M, Çelen Ş.
- Role of first trimester inflammatory markers in predicting adverse obstetric and neonatal outcomes in maternal obesity 1124-1128**
Akgün Aktaş B, Gençslan E, Başaran E, et al.
- Visual recovery after phacoemulsification surgery: assessing the role of age and gender... 1129-1132**
Türkoğlu C.
- Retrospective evaluation of the demographic, psychiatric, and clinical characteristics of patients admitted to the intensive care unit following a suicide attempt 1133-1142**
Şimşek MH, Korkmaz U.
- An underrecognized aspect of oncology care: temporomandibular joint evaluation in patients with head and neck cancers 1143-1148**
Öçbe M, Medişoğlu MS, Öçbe ÇB.
- Relationship between the success of arthrocentesis and preoperative magnetic resonance imaging findings 1149-1154**
Yüceer Çetiner E, Kazan D, Baş Akkor B.
- Feasibility of androgen receptor pathway inhibitors in prostate cancer: impact of comorbidities, age, and concomitant medication use-a multicenter cohort study 1155-1163**
İlhan N, Doğan A, Çubukçu E, et al.
- Evaluation of distribution and revisit rates of pediatric otorhinolaryngologic diseases visits to emergency department and outpatient clinics 1164-1170**
Yüksel Y, Gür ÖE.
- Comparison of operative and nonoperative treatment of midshaft clavicle fractures: early superiority of surgery with comparable long-term outcomes..... 1171-1175**
Güneş Z, Erdoğan E.
- Determination of reproductive health attitudes and affecting factors of women aged 15-49 1176-1181**
Utkualp N, Başoğlu M.
- Pregnancy outcomes in women with Behçet's disease 1182-1186**
Muluk E.

Is joint involvement in systemic lupus erythematosus always inflammatory?

 Zehra Özsoy,  Buğu Bulat,  Güllü Sandal Uzun,  Mustafa Ekici,  Şerife Asya Germe,  Levent Kılıç,  Ali İhsan Ertemli,  Ali Akdoğan

Division of Rheumatology, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Cite this article as: Özsoy Z, Bulat B, Sandal Uzun G, et al. Is joint involvement in systemic lupus erythematosus always inflammatory? *J Health Sci Med.* 2025;8(6):953-958.

Received: 21.08.2025

Accepted: 07.09.2025

Published: 25.10.2025

ABSTRACT

Aims: Systemic lupus erythematosus (SLE) is a chronic autoimmune connective tissue disease that can affect multiple organs and systems. Arthralgia and myalgia are common symptoms in patients with SLE, and their etiology may be inflammatory or may result from joint hypermobility or fibromyalgia. This study aimed to investigate whether musculoskeletal complaints in SLE patients are attributable solely to inflammatory activity or whether they may also be associated with hypermobility or fibromyalgia.

Methods: Patients who fulfilled the 2012 SLE classification criteria were included. Patients who had undergone surgical procedures affecting joint mobility within the last 6 months, or those with concomitant diseases such as rheumatoid arthritis or inflammatory myositis that may present with inflammatory arthritis, were excluded from the study. Disease activity at the time of examination was assessed using the SLE Disease Activity Index-2000 (SLEDAI-2K). Hypermobility was evaluated using the Beighton Hypermobility Score, and fibromyalgia was assessed according to the 2016 Fibromyalgia Diagnostic Criteria. Patients were divided into three groups: those with hypermobility, those with fibromyalgia, and those with neither condition. Demographic characteristics, comorbidities, medications, disease activity, and pain scores were analyzed.

Results: Six patients with both fibromyalgia and hypermobility were excluded. A total of 120 patients were analyzed, of whom 104 (86.7%) were female. The median (min-max) disease duration was 12 (1-38) years. Hypermobility was detected in 25 patients (20.8%) and fibromyalgia in 28 patients (23.3%). Overall, 44.1% of the cohort had either hypermobility or fibromyalgia. No significant differences were found between groups in terms of comorbidities. Patients with fibromyalgia had higher median patient VAS and pain scores compared with the other two groups, while their SLEDAI-2K activity scores were lower. Arthritis, according to SLEDAI-2K, was more frequent in the group without hypermobility or fibromyalgia. A significant difference in pulse-steroid use was found between the hypermobility and fibromyalgia groups ($p=0.01$) and between the fibromyalgia and neither group ($p=0.02$).

Conclusion: Severe musculoskeletal pain due to hypermobility may mimic arthralgia or arthritis and may be misinterpreted by clinicians as disease progression. This could lead to unnecessary immunosuppressive therapy (risk of overtreatment), which could lead to increased risk of infection and liver and kidney dysfunction. Therefore, hypermobility and fibromyalgia, which may coexist in the course of SLE, should be thoroughly evaluated. In our study, a significant frequency (almost half of the patients) of pain was due to non-inflammatory causes.

Keywords: Systemic lupus erythematosus, inflammation, hypermobility, fibromyalgia, arthritis

INTRODUCTION

Systemic lupus erythematosus (SLE) is a chronic autoimmune connective tissue disease that can affect multiple organs and systems. Genetic, environmental, and hormonal factors play a role in its etiology.¹ The prevalence of joint involvement in patients with SLE ranges from 69% to 95%. Unlike inflammatory arthritides such as rheumatoid arthritis, joint involvement in SLE typically follows a non-erosive course. Limitation of range of motion is uncommon; however, subluxations and laxity may occur.² Joint hypermobility is defined as the extension of joint range of motion beyond the normal limits established for

age, sex, and ethnicity during passive movement. It may occur as an isolated condition or in association with rheumatologic diseases. Benign joint hypermobility syndrome (BJHS) is characterized by widespread joint laxity and musculoskeletal pain in the absence of an associated rheumatologic disease.³

The most widely used method for assessing hypermobility is the five-item criterion set developed by Beighton et al.⁴ Each extremity is evaluated separately, and the patient receives a score ranging from 0 to 9. A score of ≥ 4 is considered indicative

Corresponding Author: Zehra Özsoy, dr.zehraduman@hotmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

of joint hypermobility. Arthralgia and myalgia are common symptoms in both SLE and BJHS. Joint hypermobility has been reported in approximately 25% of SLE patients in the literature.⁵ Another condition characterized by widespread musculoskeletal pain is fibromyalgia. The diagnosis of fibromyalgia requires fulfillment of all the items outlined in the 2016 American College of Rheumatology (ACR) diagnostic criteria.⁶ The prevalence of fibromyalgia in SLE patients has been reported to reach as high as 60% in some studies.⁷

Arthralgia and myalgia are frequent symptoms in SLE patients. While these symptoms may arise from SLE disease activity, they may also be due to underlying joint hypermobility and/or fibromyalgia. This study aimed to determine the prevalence of hypermobility and fibromyalgia in patients with SLE and to investigate their impact on disease activity.

METHODS

This study was conducted in accordance with the 2013 revision of the Declaration of Helsinki and was approved by the Hacettepe University Health Sciences Researches Ethics Committee (Date: 03.09.2024, Decision No: 2024/15-41).

This study included 120 patients followed at the Hacettepe University Department of Rheumatology between 10 September 2024, and 15 June 2025, who fulfilled the 2012 Systemic Lupus International Collaborating Clinics (SLICC) classification criteria. Patients who had undergone surgical procedures affecting joint mobility in the previous 6 months or who had comorbid diagnoses that could present with inflammatory arthritis, such as rheumatoid arthritis or inflammatory myositis, were excluded. Our study was conducted as a prospective cohort.

Hypermobility was assessed in SLE patients using the Beighton Hypermobility Score,⁴ and fibromyalgia was defined according to the ACR 2016 fibromyalgia diagnostic criteria.⁶ Disease activity at the time of examination was determined using the SLE Disease Activity Index-2000 (SLEDAI-2K) assessment form.⁸ Hypermobility, fibromyalgia, and disease activity were assessed by the same rheumatologist at the same visit. Clinical and laboratory data of the included patients were retrieved from the hospital automation system and patient files.

Six patients with both fibromyalgia and hypermobility were excluded from the study and not included in the analysis to avoid diagnostic overlap. Patients were divided into three groups: those with hypermobility, those with fibromyalgia, and those with neither condition. Demographic characteristics (age, sex), presenting symptoms, comorbidities (hypertension, diabetes mellitus, coronary artery disease, chronic kidney disease, COPD, asthma), medications (mycophenolate mofetil, azathioprine, methotrexate, leflunomide, cyclosporine, cyclophosphamide, rituximab, IVIG, glucocorticoids, hydroxychloroquine), SLEDAI-2K activity scores, and pain assessments (pain score and VAS) were compared.

Statistical Analysis

Patient data were analyzed using the IBM SPSS Statistics version 23 biostatistics software. The normality of distribution

for numerical variables was assessed both visually (histograms and probability plots) and analytically (Kolmogorov-Smirnov/Shapiro-Wilk tests). Descriptive statistics were presented as mean±standard deviation for normally distributed variables, and as median and interquartile range for non-normally distributed variables. Likewise, categorical and non-numerical variables were presented as frequency table. For comparisons of means between two independent groups with normal distribution, the student's T test was used, whereas for non-normally distributed data the Mann-Whitney U test was applied. For dependent two-group comparisons, the paired sample t-test and Wilcoxon signed-rank test were used for normally and non-normally distributed variables, respectively. In analyses involving more than two groups, one-way ANOVA or Kruskal-Wallis test was performed for independent groups depending on distributional assumptions, while repeated measures ANOVA or the Friedman test was applied for dependent groups. For categorical variables or frequency comparisons, the Chi-square or Fisher's exact test was employed, depending on sample distribution. A p-value <0.05 was considered statistically significant. The effects of hypermobility and fibromyalgia on disease activity were analyzed by logistic regression.

RESULTS

Baseline Characteristics

At baseline, six patients with both fibromyalgia and hypermobility were excluded, and the remaining 120 patients were included in the analysis, of whom 104 (86.7%) were female. The median (min-max) age was 40 (18-77) years. The median (min-max) disease duration was 12 (1-38) years. Hypermobility was detected in 25 patients (20.8%) and fibromyalgia in 28 patients (23.3%). Overall, 44.1% of the cohort had either hypermobility or fibromyalgia.

Clinical Outcomes

Patients with hypermobility were younger compared with those with fibromyalgia or those without either condition. Patients with fibromyalgia had higher median patient VAS and pain scores compared with the other two groups, while their SLEDAI-2K activity scores were lower ([Table 1](#)). Arthritis, according to SLEDAI-2K, was more frequent in the group without hypermobility or fibromyalgia. There was no difference between the groups in terms of organ/system involvement at the time of study inclusion. However, hypocomplementemia were more common in the hypermobility group. In addition, CRP levels were higher in the group without hypermobility or fibromyalgia ([Table 1](#)). No difference was found between the groups in terms of comorbidities ([Table 2](#)). When immunosuppressive treatments received at any time were analyzed, there was no difference between groups except for history of pulse-steroid therapy (p=0.02) ([Table 2](#)). Pairwise comparisons of pulse-steroid history revealed significant differences between the hypermobility and fibromyalgia groups (p=0.01) and between the fibromyalgia and neither group (p=0.02).

In baseline analyses, hypermobility was present in 24.3% of patients with active disease (SLEDAI ≥3) compared to 16.0%

Table 1. Comparison of patient groups with hypermobility only, fibromyalgia only, and neither

	Hypermobility n (%) 25 (20.8)	Those with fibromyalgia n (%) 28 (23.3)	Without hypermobility and fibromyalgia n (%): 67 (55.9)	p
Female	23 (92)	27 (96.4)	54 (80.6)	0.08
Age	27 (19-58)	49.5 (18-77)	41 (18-73)	<0.001
Disease duration (years)	9 (2-26)	13 (1-31)	11 (1-38)	0.34
Patient VAS	20 (0-90)	40 (20-80)	20 (0-100)	<0.001
Pain assessment	0 (0-3)	2 (0-4)	0 (0-4)	<0.001
SLEDAI-2K activity score	4 (0-35)	2 (0-12)	4 (0-36)	0.02
The organ/system involvements at the time of inclusion n (%)				
Arthritis	5 (20)	1 (3.6)	17 (25.4)	0.04
Fever	9 (36)	13 (46.4)	32 (47.8)	0.59
Hematopoietic system involvement	16 (64)	11 (39.3)	33 (49.3)	0.19
Central nervous system involvement	4 (16)	2 (7.1)	12 (17.9)	0.4
Skin involvement	21 (84)	23 (82.1)	50 (74.6)	0.53
Renal involvement	11 (44)	7 (25)	27 (40.3)	0.28
Cardiac involvement	3 (12)	2 (7.1)	9 (13.4)	0.68
Pleural effusion	3 (12)	1 (3.6)	10 (14.9)	0.29
Laboratory findings at the time of inclusion				
Anti-ds DNA positivity	20 (80)	21 (75)	58 (86.6)	0.37
Complement deficiency	24 (96)	20 (71.4)	57 (85.1)	0.04
CRP levels, mg/dl	2 (1-7)	2 (1-4)	8 (1-26)	<0.001
AFAS antibody positivity	15 (60)	10 (37)	29 (45.3)	0.24
RF positivity	5 (27.8)	3 (12)	10 (16.9)	0.39
CCP positivity	0 (0)	2 (12.5)	2 (4.1)	0.27

When the three groups were compared, the significant parameters were compared in paired groups. There was a difference in median age between the group with hypermobility and fibromyalgia ($p<0.001$), between the group with hypermobility and neither group ($p<0.001$), and between the group with fibromyalgia and neither group ($p: 0.028$). There was a difference in the median patient VAS value between the groups with hypermobility and fibromyalgia ($p: 0.003$), and between the groups with fibromyalgia and neither ($p<0.001$). There was a difference in pain assessment between the groups with hypermobility and fibromyalgia ($p: 0.002$), between the groups with fibromyalgia and neither ($p<0.001$). There was a difference in the median SLEDAI-2K Activity Score between the groups with hypermobility and fibromyalgia ($p: 0.01$), and between the groups with fibromyalgia and neither ($p: 0.01$). There was a difference in the frequency of arthritis according to the SLEDAI-2K Activity Score between the groups with fibromyalgia and neither ($p<0.01$). There was a difference in hypocomplementemia between the hypermobility and fibromyalgia groups ($p: 0.01$). There was a difference in the median patient CRP value between the groups with hypermobility and neither ($p<0.001$), and between the groups with fibromyalgia and neither ($p<0.001$).

VAS: Visual Analog Scale, SLEDAI-2K: SLE Disease Activity Index-2000, DNA: Deoxyribose nucleic acid, CRP: C-reactive protein, AFAS: Antiphospholipid syndrome, RF: Rheumatoid factor, CCP: Cyclic citrullinated peptide

Table 2. Comparison of comorbidities and treatments according to patient groups with hypermobility only, fibromyalgia only, and neither

		Hypermobility n (%) 25 (20.8)	Those with fibromyalgia n (%) 28 (23.3)	Without hypermobility and fibromyalgia n (%): 67 (55.9)	p
Comorbidities	Hypertension	1 (4)	8 (28.6)	13 (19.4)	0.06
	Diabetes mellitus	0 (0)	3 (10.7)	2 (3)	0.11
	Coronary artery disease	0 (0)	1 (3.6)	4 (6)	0.43
	Chronic renal failure	5 (20)	5 (17.9)	6 (9)	0.27
	Antiphospholipid antibody syndrome	3 (12)	3 (10.7)	5 (7.5)	0.75
Treatments	Mycophenolate mofetil	11 (44)	11 (39.3)	31 (46.2)	0.82
	Azathioprine	13 (52)	14 (50)	26 (38.8)	0.4
	Cyclosporine	2 (8)	1 (3.6)	3 (4.5)	0.72
	Cyclophosphamide	6 (24)	3 (10.7)	21 (31.3)	0.1
	Rituximab	3 (12)	5 (17.9)	14 (21.2)	0.59
	Methotrexate	1 (4)	0 (0)	7 (10.4)	0.14
	Leflunomide	0 (0)	2 (7.1)	3 (4.5)	0.42
	IVIg	5 (20)	2 (7.1)	14 (20.9)	0.25
	Daily corticosteroid use	21 (84)	25 (89.3)	56 (83.6)	0.76
	History of pulse-steroid therapy	16 (64)	8 (28.6)	36 (53.7)	0.02

Pairwise comparisons of pulse-steroid history revealed significant differences between the hypermobility and fibromyalgia groups ($p=0.01$) and between the fibromyalgia and neither group ($p=0.02$).

IVIg: Intravenous immunoglobulin

of those without activity ($p=0.271$), indicating no significant difference. In contrast, fibromyalgia was observed more frequently among patients without active disease (34.0% vs. 15.7%, $p=0.020$). Consistently, univariate logistic regression demonstrated that fibromyalgia was associated with a lower likelihood of active disease (OR=0.362, 95% CI 0.152-0.864, $p=0.022$), whereas hypermobility showed no significant association (OR=1.684, 95% CI 0.663-4.280, $p=0.274$).

DISCUSSION

In our study, the majority of patients were female. Hypermobility was present in 20.8% of patients, and fibromyalgia in 23.3%. Patients with hypermobility were younger compared with the other groups. Patients with fibromyalgia had higher patient VAS and pain scores compared with the other two groups, while their SLEDAI-2K activity scores were lower. According to the SLEDAI-2K, arthritis was more frequent in patients without hypermobility or fibromyalgia, whereas, hypocomplementemia were more common in the hypermobility group. A history of pulse-steroid therapy was less frequent among patients with fibromyalgia.

Joint involvement is one of the most common manifestations in patients with SLE and may occur in up to 90% of cases. According to the 2012 SLICC criteria, joint involvement in SLE is defined as “synovitis involving two or more joints, characterized by swelling or effusion, or tenderness in two or more joints with morning stiffness lasting at least 30 minutes”.⁸ Although not life-threatening, joint involvement significantly impairs quality of life, leading to functional limitations and even disability. The frequency of work loss in SLE patients during the first 5 years of disease onset ranges between 15% and 40%.^{9,10} Joint involvement may occur at any stage of the disease. It is characterized by wide heterogeneity in phenotype and severity, ranging from minor arthralgia to erosive arthritis.^{11,12} Arthralgia most frequently involves the hand joints and may sometimes be associated with morning stiffness.¹¹ In a study of 1000 patients, Cervera et al.¹³ reported arthritis in 84% of the cohort. Similarly, in a more recent lupus cohort, musculoskeletal involvement was reported in 80% of patients.¹⁴ In our study, arthritis was present in 19.2% according to the SLEDAI-2K, and was more frequent in the group without hypermobility or fibromyalgia. SLE-related arthritis is typically characterized by symmetric involvement of small joints. In most patients, bone erosions are absent.¹¹ In long-standing disease, however, severe joint laxity and deformities may develop, a condition referred to as Jaccoud’s arthropathy (JA).¹⁵ JA affects approximately 3-13% of patients and may result in joint subluxation.^{16,17} When evaluating patients with SLE, rheumatologists should consider the possibility of hypermobility and fibromyalgia in the presence of arthralgia without specific signs of arthritis, such as joint swelling, redness, or warmth. Systematic screening for fibromyalgia and hypermobility should be performed in SLE patients with persistent joint pain.

The gold standard for assessing SLE disease activity has not yet been established. An ideal activity index should reflect both improvement and deterioration across the different systems involved during the disease course, and should distinguish

activity from chronic damage. Among the available indices, SLEDAI, revised in 2002, is the most commonly used in observational and randomized controlled studies. It consists of 24 objective variables, including 16 clinical and 8 laboratory items, related to disease activity within the previous 28 days.^{18,19} However, this tool may not capture all possible features of joint involvement in SLE patients, such as inflammatory arthralgia or monoarthritis. In our cohort, disease activity was assessed using the SLEDAI-2K, and when groups were compared, it was revealed that patients with fibromyalgia had lower disease activity scores. In addition, univariate logistic regression demonstrated that fibromyalgia was associated with a lower likelihood of active disease. Overall, there is a strong scientific consensus that FM does not directly increase SLE activity but negatively impacts symptom burden and quality of life.²⁰ While there is no definitive explanation for this, patients with higher disease activity likely experience symptoms related to disease burden. Additionally, it may be difficult to diagnose fibromyalgia in someone with active disease due to the disease burden. Considering that fibromyalgia patients focus more on their perception of pain and therefore make more hospital admissions, these patients may have lower disease activity because they are under more doctor’s care.

Joint laxity in rheumatic diseases has been investigated in various studies over the years. Benign joint hypermobility syndrome (BJHS) has been reported in 5-57% of young women and in 2-35% of men in the general population.²¹ Although BEHS historically considered to be benign, one study reported that the percentage of patients with BEHS who had at least one rheumatologic disease was 67.1% in those who were medically examined, 33.3% in those with limited examination, and 9.2% in those who were not examined at all, when comprehensive serologic and radiographic studies were performed for musculoskeletal complaints in addition to physical examination (MAK6- 11). The five most common rheumatologic diseases in BJHS patients were fibromyalgia (22 cases), psoriasis (22 cases), ankylosing spondylitis (11 cases), psoriatic arthritis (11 cases), and rheumatoid arthritis (9 cases).²² These findings highlight the need for thorough clinical assessment in BJHS patients to ensure timely diagnosis of coexisting rheumatologic diseases.

In a study examining 378 patients referred to a rheumatology clinic, hypermobility was identified in 50 patients (13.2%). Compared to controls without hypermobility, the most frequent clinical diagnosis in hypermobile patients was soft tissue rheumatism, observed in 67% versus 25% ($p<0.001$). Fibromyalgia syndrome was present in 30% of hypermobile patients compared to 8% of non-hypermobile patients ($p<0.001$), whereas the least common diagnosis was inflammatory arthritis, observed in 4% versus 32% ($p<0.001$).²³ These results suggest a relationship between hypermobility and soft tissue rheumatic complaints, which may be clinically useful for rheumatologists. In our study, six patients with both fibromyalgia and hypermobility were excluded from the study and not included in the analysis to avoid diagnostic overlap.

Although there are few studies on the relationship between SLE and hypermobility in the literature, it has been shown that hypermobility is associated with inflammatory

arthropathies, particularly SLE.²⁴ In one study, 10 of 59 patients with SLE (17%) were found to be hypermobile.²⁵ Studies conducted in Iraq in 2018 and in Türkiye in 2024 reported a hypermobility prevalence of 52.8% and 50%, respectively.^{26,27} Another study conducted with 81 patients found a statistically significant difference in hypermobility between SLE patients and healthy populations.²⁸ whereas Bleifeld and Inglis²⁹ described localized hand hypermobility in half of SLE patients. In addition to Bridges et al.'s²⁴ findings that 25% of SLE patients are hypermobile, another study with 16 SLE patients and 19 healthy controls, as well as Klemp et al.'s³⁰ study, found hypermobility to be more common in SLE; however, no statistically significant difference between groups was reported.⁵ In our cohort, 20.8% of patients exhibited hypermobility.

As expected from the literature, rheumatology patients are predominantly female (66.9% female, 33.1% male), and the BEHS patient population was also predominantly female (84.4%).²² A similar female-to-male ratio (89% female, 11% male) has been reported in an Italian BEHS cohort.³¹ Differences in muscle structure, sex hormones, and pain perception between males and females have been suggested, although genetic causes remain largely unknown,^{31,32} and further genetic studies are needed to determine whether it is an X-linked condition. Consistent with the literature, 92% of patients with hypermobility in our study were female. Interestingly, in a controlled study of SLE patients with severe disease requiring long-term corticosteroid therapy, hypermobility was not associated with age at disease onset, disease duration, or corticosteroid treatment.³⁰ In our study, 64% of hypermobile patients had a history of pulse steroid therapy, while disease duration was similar across all three groups. Wallace et al.³³ reported a fibromyalgia prevalence of 22% in 464 SLE patients, and Morand et al.³⁴ reported 25.3% in a cohort of 87 SLE patients. Another study found fibromyalgia prevalence in SLE to be as high as 61%.³⁵ In our cohort, consistent with the literature, the prevalence of fibromyalgia was 23.3%.

Limitations

The absence of a control group in our study is a significant limitation. Multicenter studies including control groups are needed to better elucidate the relationship between hypermobility and lupus and to provide guidance in clinical practice.

CONCLUSION

There are insufficient studies in the literature regarding our cohort. Although our study's primary strength is its single-center presence, we contribute to the literature with our substantial patient population. We believe that clinicians should be aware of the importance of considering the diagnoses of hypermobility and fibromyalgia, in addition to disease activation, when managing arthralgia in SLE patients. This will prevent unnecessary immunosuppressive therapy, which can result in complications such as unnecessary infection and organ dysfunction, and which also carries additional financial costs. In conclusion, musculoskeletal symptoms are common in most SLE patients. While many of these symptoms are related to disease activity, some may

be associated with hypermobility or fibromyalgia. Detecting these conditions is essential for disease management. The high prevalence of joint involvement in SLE patients and the clinical heterogeneity require increased awareness and thorough evaluation. Hypermobility disorder can cause severe musculoskeletal pain, which may be mistaken for disease-related arthralgia or arthritis in SLE and perceived by clinicians as disease progression, potentially leading to unnecessary immunosuppressive therapy or escalation of treatment. Therefore, hypermobility and fibromyalgia, which may coexist in the course of SLE, should be thoroughly evaluated. Although arthralgia is a complaint shared across different diseases, applying diagnostic criteria minimizes uncertainty in diagnosis. Distinctions should be made between complaints associated with fibromyalgia and HEDS. However, any clinical diagnosis is subject to inter-clinician variability, highlighting the need for further clinical studies.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Hacettepe University Health Sciences Researches Ethics Committee (Date: 03.09.2024, Decision No: 2024/15-41).

Informed Consent

All patients signed and free and informed consent form.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Data Sharing Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

1. Mills JA. Systemic lupus erythematosus. *N Engl J Med*. 1994;330(26):1871-1879. doi:10.1056/NEJM199406303302608
2. Ceccarelli F, Perricone C, Cipriano E, et al. Joint involvement in systemic lupus erythematosus: from pathogenesis to clinical assessment. *Semin Arthritis Rheum*. 2017;47(1):53-64. doi:10.1016/j.semarthrit.2017.03.022
3. Boudreau PA, Steiman I, Mior S. Clinical management of benign joint hypermobility syndrome: a case series. *J Can Chiropr Assoc*. 2020;64(1):43-54.
4. Beighton P, Solomon L, Soskolne CL. Articular mobility in an African population. *Ann Rheum Dis*. 1973;32:413-418. doi:10.1136/ard.32.5.413
5. Johnsson P, Kohlin C, Rasmussen G, Bengtsson A, Sturfelt G, Nived O. Differences in body structure and function between patients with systemic lupus erythematosus and healthy individuals, with particular reference to joint hypermobility. *Scand J Rheumatol*. 2016;46(1):40-43. doi:10.3109/03009742.2016.1158313

6. Wolfe F, Clauw DJ, Fitzcharles MA, et al. 2016 revisions to the 2010/2011 fibromyalgia diagnostic criteria. *Semin Arthritis Rheum.* 2016;46(3):319-329. doi:10.1016/j.semarthrit.2016.08.012
7. Buskila D, Press J, Abu-Shakra M. Fibromyalgia in systemic lupus erythematosus: prevalence and clinical implications. *Clin Rev Allergy Immunol.* 2003;25(1):25-28. doi:10.1385/CRIAI:25:1:25
8. Petri M, Orbai AM, Alarcon GS, et al. Derivation and validation of the Systemic Lupus International Collaborating Clinics classification criteria for systemic lupus erythematosus. *Arthritis Rheum.* 2012;64(8):2677-2686. doi:10.1002/art.34473
9. Bertoli AM, Fernandez M, Alarcon GS, Vila LM, Reveille JD. Systemic lupus erythematosus in a multiethnic US cohort LUMINA (XLI): factors predictive of self-reported work disability. *Ann Rheum Dis.* 2007;66(1):12-17. doi:10.1136/ard.2006.055343
10. Al Dhanhani AM, Gignac MA, Su J, Fortin PR. Work disability in systemic lupus erythematosus. *Arthritis Rheum.* 2009;61(3):378-385. doi:10.1002/art.24347
11. Ball EM, Bell AL. Lupus arthritis-do we have a clinically useful classification? *Rheumatology (Oxford).* 2012;51(5):771-779. doi:10.1093/rheumatology/ker381
12. Gormezano NW, Silva CA, Aikawa NE, et al. Chronic arthritis in systemic lupus erythematosus: distinct features in 336 paediatric and 1830 adult patients. *Clin Rheumatol.* 2016;35(1):227-231. doi:10.1007/s10067-015-3127-3
13. Cervera R, Khamashta MA, Font J, et al. Systemic lupus erythematosus: clinical and immunologic patterns of disease expression in a cohort of 1,000 patients. The European Working Party on Systemic Lupus Erythematosus. *Medicine (Baltimore).* 1993;72(2):113-124.
14. Conti F, Ceccarelli F, Perricone C, et al. The chronic damage in systemic lupus erythematosus is driven by flares, glucocorticoids and antiphospholipid antibodies: results from a monocentric cohort. *Lupus.* 2016;25(7):719-726. doi:10.1177/0961203315627199
15. Chattopadhyay A, Ghosh A. Jaccoud's arthropathy. *Rheumatology (Oxford).* 2020;59(10):p3117. doi:10.1093/rheumatology/keaa128
16. Mediawake R, Isenberg DA, Schellekens GA, van Venrooij WJ. Use of anti-citrullinated peptide and anti-RA33 antibodies in distinguishing erosive arthritis in patients with systemic lupus erythematosus and rheumatoid arthritis. *Ann Rheum Dis.* 2001;60(1):67-68. doi:10.1136/ard.60.1.67
17. Van Vugt RM, Derksen RH, Kater L, Bijlsma JW. Deforming arthropathy or lupus and rhuus hands in systemic lupus erythematosus. *Ann Rheum Dis.* 1998;57(9):540-544. doi:10.1136/ard.57.9.540
18. Gladman DD, Ibanez D, Urowitz MB. Systemic lupus erythematosus disease Activity Index 2000. *J Rheumatol.* 2002;29(2):288-291.
19. Gladman DD, Goldsmith CH, Urowitz MB, et al. Crosscultural validation and reliability of 3 disease activity indices in systemic lupus erythematosus. *J Rheumatol.* 1992;19(4):608-611.
20. Wolfe F, Petri M, Alarcón GS, et al. Fibromyalgia, systemic lupus erythematosus, and evaluation of SLE activity. *J Rheumatol.* 2009;36(1):82-88. doi:10.3899/jrheum.080212
21. Morlino S, Castori M. Placing joint hypermobility in context: traits, disorders and syndromes. *Br Med Bull.* 2023;147(1):90-107. doi:10.1093/bmb/ldad013
22. Rodgers KR, Gui J, Dinulos MB, Chou RC. Ehlers-Danlos syndrome hypermobility type is associated with rheumatic diseases. *Sci Rep.* 2017;7:39636. doi:10.1038/srep39636
23. Hudson N, Starr MR, Esdaile JM, Fitzcharles MA. Diagnostic associations with hypermobility in rheumatology patients. *Br J Rheumatol.* 1995;34(12):1157-1161. doi:10.1093/rheumatology/34.12.1157
24. Bridges AJ, Smith E, Reid J. Joint hypermobility in adults referred to rheumatology clinics. *Ann Rheum Dis.* 1992;51(6):793-796. doi:10.1136/ard.51.6.793
25. Babini SM, Cocco JAM, Babini JC, de la Sota M, Arturi A, Marcos JC. Atlantoaxial subluxation in systemic lupus erythematosus: further evidence of tendinous alterations. *J Rheumatol.* 1990;17(2):173-177.
26. Mercan R, Bulut Gokten D, Yümün Kavak F. Frequency of hypermobility in patients diagnosed with systemic lupus erythematosus and demographic, clinical characteristics of these patients. *Eskisehir Med J.* 2024;5(1):19-22. doi:10.48176/esmj.2024.154
27. Al-Osami MH, Mohammed S, Ali WI, Fawzi HA. The relationship between hypermobility syndrome and systemic lupus erythematosus. *Ind J Public Health Res Develop.* 2018;9(12):526-532.
28. Gumà M, Olivé A, Roca J, et al. Association of systemic lupus erythematosus and hypermobility. *Ann Rheum Dis.* 2002;61(11):1024-1026. doi:10.1136/ard.61.11.1024
29. Bleifeld CJ, Inglis AE. The hand in systemic lupus erythematosus. *J Bone Joint Surg Am.* 1974;56(6):1207-1215.
30. Klemp P, Majoo FL, Charlton D. Articular mobility in systemic lupus erythematosus (SLE). *Clin Rheumatol.* 1987;6(2):202-207. doi:10.1007/BF02201025
31. Castori M. Ehlers-Danlos syndrome, hypermobility type: an underdiagnosed hereditary connective tissue disorder with mucocutaneous, articular, and systemic manifestations. *ISRN Dermatol.* 2012;2012:751768. doi:10.5402/2012/751768
32. Castori M, Camerota F, Celletti C, Grammatico P, Padua L. Ehlers-Danlos syndrome hypermobility type and the excess of affected females: possible mechanisms and perspectives. *Am J Med Genet A.* 2010;152A(9):2406-2408. doi:10.1002/ajmg.a.33585
33. Wallace DJ, Schwartz E, Chi-Lin H, Peter JB. The 'rule out lupus' rheumatology consultation: clinical outcomes and perspectives. *J Clin Rheumatol.* 1995;1(3):158-164.
34. Morand EF, Miller MH, Whittingham S, Littlejohn GO. Fibromyalgia syndrome and disease activity in systemic lupus erythematosus. *Lupus.* 1994;3(3):187-191. doi:10.1177/096120339400300310
35. Romano TJ. Management of pain in arthritis. *Am J Pain Manag.* 1992;2(3):211-214.

Investigation of systemic treatment responses and prognostic factors in metastatic gastroenteropancreatic neuroendocrine carcinomas (GEP-NECs)

Alper Türkel¹, Duygu Ercan Uzundal², Fatih Gürler², İlknur Deliktaş Onur³,
Serkan Gülcü⁴, Ülkü Yalçıntaş Arslan³

¹Department of Medical Oncology, Eskişehir City Hospital, Eskişehir, Türkiye

²Division of Medical Oncology, Department of Internal Medicine, Faculty of Medicine, Gazi University, Ankara, Türkiye

³Department of Medical Oncology, Dr. Abdurrahman Yurtaslan Ankara Oncology Training and Research Hospital, University of Health Sciences, Ankara, Türkiye

⁴Department of Medical Oncology, Ankara Etlik City Hospital, Ankara, Türkiye

Cite this article as: Türkel A, Ercan Uzundal D, Gürler F, Deliktaş Onur İ, Gülcü S, Yalçıntaş Arslan Ü. Investigation of systemic treatment responses and prognostic factors in metastatic gastroenteropancreatic neuroendocrine carcinomas (GEP-NECs). *J Health Sci Med.* 2025;8(6):959-965.

Received: 17.07.2025

Accepted: 13.09.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to evaluate responses to first- and second-line treatments, survival outcomes, and potential factors affecting survival in patients with metastatic gastroenteropancreatic neuroendocrine carcinoma (GEP-NEC).

Methods: This retrospective study included 42 patients with metastatic GEP-NEC from three centers. First and second-line chemotherapy regimens were examined. The median progression-free survival (mPFS1 and mPFS2) for first and second-line treatments and overall survival (mOS) were evaluated. The effects of CRP levels, LDH levels, neutrophil-to-lymphocyte ratio (NLR), and platelet-to-lymphocyte ratio (PLR) on survival were investigated as potential prognostic factors.

Results: Median follow-up was 15 months (3-113). mPFS1 was 5 months (95% CI: 2.88-7.11) and mPFS2 was 2 months (95% CI: 1.67-2.32). The mOS was 15 months (95% CI: 10.90-19.09). Curative surgery patients had a 28-month mOS, and patients who could not undergo surgery had a mOS of 14 months ($p=0.010$). Patients with $NLR < 2.56$ had a longer mOS (22 months vs. 14 months, $p=0.019$) than those with $NLR \geq 2.56$. Multivariate analysis showed that lower NLR scores and Ki67 values were associated with longer mOS [HR: 0.38 (95% CI: 0.18-0.79), $p=0.010$ and HR: 0.41 (95% CI: 0.20-0.84), $p=0.015$]

Conclusion: Our study showed that higher NLR scores and Ki67 values significantly worsened mOS. Survival outcomes (OS, PFS1, and PFS2) aligned with the literature. NECs need new treatments and prognostic markers due to their poor prognosis, short survival times, and lack of standard guidelines.

Keywords: Extrapulmonary, neuroendocrine carcinoma, GEP-NEC, NLR

INTRODUCTION

Neuroendocrine carcinomas (NECs) are rare, high-grade, poorly differentiated, aggressive tumors with a poor prognosis.¹⁻³ Although approximately 90% of neuroendocrine carcinomas are of pulmonary origin, the most common site for extrapulmonary NECs is the gastroenteropancreatic (GEP) system. A study including 162,983 NEC cases found that only 9% of patients had extrapulmonary NECs, with 37.4% of these originating from the gastrointestinal tract.⁴ GEP-NECs are chemosensitive tumors; a combination of platinum and etoposide is the preferred chemotherapy regimen for treatment, similar to small-cell lung cancer.⁵⁻⁷ However, although most patients respond to first-line cytotoxic therapy, the response times are quite short [(median PFS: 4 months (3.4-4.6) 95% CI)].⁸ The current literature on second and later-line therapy is limited, and no established optimal treatment protocol exists. There are no randomized trials on this topic.

Additionally, a significant portion of patients cannot receive second-line therapy due to rapid clinical deterioration and/or declining ECOG performance status, and the best for these patients is supportive care.

Various studies have investigated some clinicopathological features and laboratory values that may affect the prognosis of neuroendocrine malignancies. These include Ki67 values,⁹ lactate dehydrogenase (LDH) levels,⁸ C-reactive protein (CRP) levels,¹⁰ white blood cell (WBC) and neutrophil count,¹¹ neutrophil-lymphocyte ratio (NLR),¹²⁻¹⁴ and platelet-lymphocyte ratio (PLR).¹⁵

This study investigated responses to first- and second-line treatments, survival outcomes, and potential factors affecting survival in patients with metastatic GEP-NEC.

Corresponding Author: Alper Türkel, turkelalperr@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

METHODS

The study was carried out with the permission of the Dr. Abdurrahman Yurtaslan Ankara Oncology Training and Research Hospital Non-interventional Clinical Researches Ethics Committee (Date: 19.10.2023, Decision No: 2023-10/97). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

This retrospective, multicenter study included 42 patients diagnosed with metastatic gastroenteropancreatic neuroendocrine carcinoma between December 2010 and December 2023. The inclusion criteria specified patients over 18 years of age with de novo or recurrent metastatic disease originating from the stomach, intestines, or pancreas who received at least two lines of chemotherapy and had complete access to their diagnostic and treatment records from the hospital registry system. Patients with pulmonary or unknown primary neuroendocrine carcinomas, those unable to receive chemotherapy, or those with incomplete diagnostic or treatment histories were excluded.

Data collected included demographic details such as age, sex, comorbidities, and smoking history, as well as clinicopathological features like tumor location, histological subtypes, Ki67 levels, sites of metastasis, first- and second-line chemotherapy regimens and cycles, and laboratory values, including WBC counts ($10^9/L$), neutrophil counts ($10^9/L$), CRP levels, LDH levels, platelet counts ($10^9/L$), NLR, and PLR.

In first-line treatment, platinum (cisplatin or carboplatin) plus etoposide treatment was preferred in the majority of patients according to treatment guidelines (although it was observed that different treatments were also preferred due to platinum ineligibility or low ECOG PS status), while in second-line, treatment selection was made according to the chemotherapy regimens which preferred more frequently by the centers where treatment decisions were made.

The patients' diagnosis, progression, and exitus or last follow-up dates were recorded. The time from initial diagnosis to progression under first-line treatment was defined as progression-free survival 1 (PFS1). The time from the date of progression after first-line therapy to progression under second-line treatment was defined as progression-free survival 2 (PFS2). The time from the date of initial diagnosis to the date of exitus or last follow-up was defined as overall survival (OS).

Statistical Analysis

Descriptive statistics were used to evaluate demographic and clinicopathological features. Survival analyses of median PFS1, median PFS2, and median OS were performed using the Kaplan-Meier method, and possible prognostic factors were compared using the log-rank test. OS was analyzed according to the time from the date of progression to the metastatic stage in patients to the date of last follow-up or exitus. The Cox regression method was used for multivariate analysis.

A p-value less than 0.05 was considered statistically significant. All statistical analyses were performed using the Statistical Package for Social Sciences (SPSS) for personal computers, version 21.0.

RESULTS

Baseline Characteristics of Patients

Of the 42 patients included, 24 (57.1%) were male and 18 (42.9%) were female. The mean age of the patients was 57.7 (± 11.1). When primary tumor localizations were examined, 17 (40.5%) patients had pancreatic, 16 (38%) patients had gastric, and 9 (21.5%) patients had colorectal origin. In pathological assessment, 12 (28.6%) patients were diagnosed with small cell carcinoma, while 30 (71.4%) patients were diagnosed with large cell carcinoma. The median Ki67 value of the patients was 60% (21%-100%). At the initial diagnosis, 13 (31%) patients were eligible for upfront surgery with curative intent. Upon examination of metastatic sites, it was observed that the two most common metastatic locations were the liver and lymph nodes. All metastatic sites of the patients, along with their baseline characteristics, are detailed in [Table 1](#).

Table 1. Baseline characteristics of patients	
Age (years), ($\pm$ SD)	57.7 (± 11.1)
Sex (M/F), n	24/18
Smoking history, n (%)	
Active smoker	13 (31)
Non-smoker	20 (47.6)
Ex-smoker	9 (21.4)
ECOG performance status	
ECOG PS 0	11 (26.2)
ECOG PS 1	30 (71.5)
ECOG PS 2	1 (2.3)
Tumor localization	
Pancreatic	17 (40.5)
Gastric	16 (38)
Colorectal	9 (21.5)
Pathological subtype	
Small-cell	12 (28.6)
Large-cell	30 (71.4)
Ki67, %	
Mean value ($\pm$ SD)	59.6 (± 29.7)
Median value	60 (21-100)
Operation with curative intent	
Yes	13 (31)
No	29 (69)
Metastases sites	
Liver	32
Lymph nodes	24
Other organs	9

SD: Standard deviation, M: Male, F: Female, ECOG: Eastern Cooperative Oncology Group, PS: Performance Scale

Treatment Regimens

All patients (n=42) had received at least two lines of chemotherapy, and all patients (n=13) who were resectable at diagnosis and underwent curative intent surgery received

adjuvant chemotherapy. Of the 13 patients mentioned, 10 received cisplatin plus etoposide, 2 received carboplatin plus etoposide, and 1 received capecitabine plus temozolomide as adjuvant therapy. In the metastatic stages, when the first-line chemotherapy regimens of the patients were examined, 33 (78.5%) patients received platinum (cisplatin or carboplatin) plus etoposide combination. The median number of treatment cycles in the first-line setting was 6 (2-19). When the second-line chemotherapy regimens were examined, 10 (23.8%) patients received topotecan, and 10 (23.8%) patients received a capecitabine plus temozolomide combination. The median number of treatment cycles in the second-line setting was 3 (1-20). The other patients were treated with different treatment regimens. In addition, the platinum plus etoposide combination was preferred in both first-line and second-line treatment in three patients because they were considered platinum-sensitive. All first-line and second-line chemotherapy regimens are detailed in [Table 2](#).

Laboratory Assessment and NLR/PLR Scoring

Possible laboratory values and scores that may affect survival were evaluated as WBC count, neutrophil count, CRP levels, LDH levels, NLR, and PLR. The median WBC count was 7.2 ($10^9/L$) (2.3-16.9), and the median neutrophil count was 4.5 ($10^9/L$) (1.4-13). The median LDH level was 220 mg/dl (60-959). The median CRP level was 5.5 mg/dl (0-115). When NLR and PLR were examined, the median NLR was 2.56 (1.09-17.8), and the median PLR was 164 (55.6-334).

Survival Analyses

The median follow-up time was 15 months (range: 3-113 months). All patients received at least two lines of chemotherapy. Five patients received 3rd-line treatment, and two patients received 4th-line treatment. Median PFS1 was 5 months (95% CI: 2.88-7.11 months). Median PFS2 was 2 months (95% CI: 1.67-2.32 months). The median OS was 15 months (95% CI: 10.90-19.09 months).

Age, sex, Ki67 value, pathological subtype, operation status, NLR, PLR, CRP, and LDH levels were examined as possible factors affecting survival. The effect of each variable on PFS1, PFS2, and OS was analyzed separately.

In the univariate analysis, the median PFS1 for patients whose primary tumor was surgically removed with curative intent

at diagnosis was 10 months, whereas the median PFS1 for patients who could not undergo surgery was 4 months; this difference was statistically significant ($p=0.010$). No significant difference was found in PFS1 for any of the other factors examined ([Table 3](#)).

In the patient group with Ki67 $\leq 60\%$, the median PFS2 was 6 months, whereas in the patient group with Ki67 $>60\%$, the median PFS2 was 2 months; this difference was statistically significant ($p=0.004$) ([Figure 1](#)). In the group with LDH ≤ 220 mg/dl, the median PFS2 was 3 months, whereas in the group with LDH >220 mg/dl, the median PFS2 was 2 months; this difference was statistically significant ($p=0.043$). No significant difference was found in PFS2 for any of the other factors examined ([Table 3](#)).

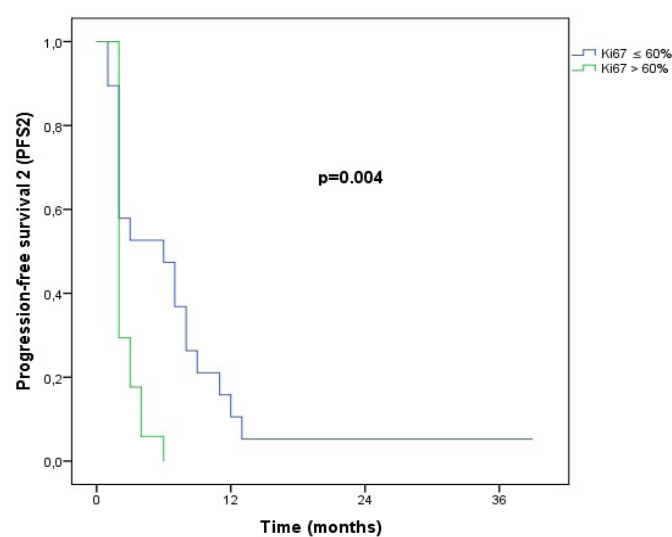


Figure 1. Kaplan-Meier curve for progression-free survival 2 according to Ki67 level

In the patient group whose primary tumor was surgically removed with curative intent, the median OS was 28 months. In contrast, the median OS was 14 months in the group that could not, and a significant difference was found between the two groups ($p=0.010$). In the patient group with NLR <2.56 , the median OS was 22 months, while in the group with NLR ≥ 2.56 , the median OS was 14 months, and a significant difference was found between the two groups ($p=0.019$) ([Figure 2](#), [Table 3](#)).

Table 2. Patients' 1st-line and 2nd-line treatment regimens

1 st -line regimens	Patients (n)	2 nd -line regimens	Patients (n)
Cisplatin plus etoposide	25	Topotecan	10
Carboplatin plus etoposide	8	Capesitabine plus temozolomide	10
Capesitabine plus temozolomide	7	Cisplatin plus etoposide	9
FOLFIRI	1	FOLFIRI	7
DCF	1	Carboplatin plus etoposide	3
Total	42	Cisplatin plus docetaxel	1
		FOLFOX	1
		Irinotecan	1
		Total	42

FOLFIRI: Folinic acid, 5-fluorouracil, and irinotecan, DCF: Docetaxel, cisplatin, and 5-fluorouracil, FOLFOX: Folinic acid, 5-fluorouracil, and oxaliplatin

Table 3. Univariate analysis of possible prognostic factors for survival

	Median PFS1 (95% CI)	p-value	Median PFS2 (95% CI)	p-value	Median OS (95% CI)	p-value
Age (years)						
≤58	4.0 (2.27-5.72)	0.248	2.0 (1.16-2.83)	0.730	15.0 (9.08-2.091)	0.456
>58	5.0 (2.80-7.19)		2.0 (1.59-2.40)		17.0 (8.23-25.76)	
Sex						
Male	5.0 (2.60-7.40)	0.785	2.0 (1.62-2.37)	0.682	14.0 (10.39-17.60)	0.440
Female	4.0 (2.61-5.38)		2.0 (0.95-3.04)		24.0 (6.48-41.51)	
Ki67 (%)						
≤60	6.0 (3.44-8.55)	0.588	6.0 (0.66-11.3)	0.004*	22.0 (8.43-35.56)	0.052
>60	5.0 (3.92-6.07)		2.0 (0.90-4.25)		13.0 (10.13-15.86)	
Pathological assessment						
Small-cell	4.0 (2.30-5.69)	0.425	2.0 (1.65-2.32)	0.611	13.0 (6.21-19.79)	0.707
Large-cell	5.0 (2.85-7.14)		2.0 (1.28-2.71)		17.0 (8.34-25.65)	
Operation at diagnosis						
Yes	10.0 (2.95-17.04)	0.010*	2.0 (1.69-2.34)	0.122	14.0 (8.87-19.12)	0.010*
No	4.0 (2.82-5.17)		2.0 (1.59-2.40)		28.0 (16.36-39.63)	
NLR						
<2.56	7.0 (4.07-9.92)	0.240	3.0 (0.38-5.61)	0.184	22.0 (6.33-37.66)	0.019*
≥2.56	4.0 (3.44-4.55)		2.0 (1.62-2.37)		14.0 (8.77-19.22)	
PLR						
<164	5.0 (2.80-7.19)	0.453	2.0 (1.55-2.44)	0.460	14.0 (7.42-20.57)	0.949
≥164	5.0 (3.07-6.92)		2.0 (1.07-2.92)		18.0 (10.38-25.61)	
CRP (mg/dl)						
≤5	5.0 (2.75-7.24)	0.907	2.0 (1.48-2.51)	0.818	13.0 (6.42-19.57)	0.661
>5	5.0 (0.51-9.48)		2.0 (1.62-2.34)		21.0 (13.97-28.02)	
LDH (mg/dl)						
≤220	6.0 (3.03-8.96)	0.789	3.0 (1.12-6.90)	0.043*	21.0 (16.15-25.84)	0.062
>220	5.0 (3.54-6.45)		2.0 (1.32-2.55)		13.0 (8.63-17.36)	
*: p<0.05, PFS1: Progression-free survival (1 st -line), PFS2: Progression-free survival (2 nd -line), OS: Overall survival, CI: Confidence interval, NLR: Neutrophil-lymphocyte ratio, CRP: C-reactive protein, LDH: Lactate dehydrogenase						

*: $p < 0.05$, PFS1: Progression-free survival (1st-line), PFS2: Progression-free survival (2nd-line), OS: Overall survival, CI: Confidence interval, NLR: Neutrophil-lymphocyte ratio, PLR: Platelet-lymphocyte ratio, CRP: C-reactive protein, LDH: Lactate dehydrogenase

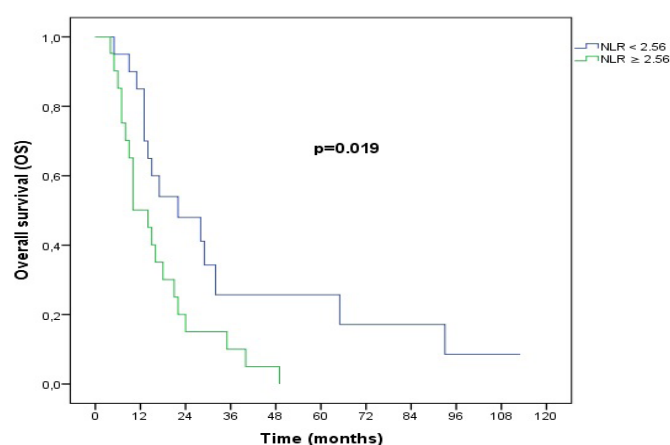


Figure 2. Kaplan-Meier curve for OS according to NLR

OS: Overall survival, NLR: Neutrophil-lymphocyte ratio

Multivariate analysis was performed to investigate the effects of curative intent surgery, Ki67 value, NLR, and LDH levels on median OS, which were identified as potential prognostic factors in the univariate analysis. As a result of the analysis, curative intent surgery and LDH levels, which were identified as significant prognostic factors in univariate analysis, lost

significance. However, a lower NLR was associated with longer median OS [HR: 0.38 (95% CI: 0.18-0.79), $p=0.010$]. Similarly, lower Ki-67 values were significantly associated with longer median OS [HR: 0.41 (95% CI: 0.20-0.84), $p=0.015$] (Table 4).

Table 4. Cox regression model for predicting the independent factors for median overall survival

	HR (95% CI)	p-value
Operation with curative intent		
No	Ref	
Yes	1.893 (0.798-4.492)	0.148
NLR		
≥2.56	Ref	
<2.56	0.388 (0.189-0.794)	0.010*
Ki67 (%)		
>60	Ref	
≤60	0.411 (0.201-0.841)	0.015*
LDH (mg/dl)		
≤220	Ref	
>220	1.048 (0.418-2.630)	0.920

HR: Hazard ratio, NLR: Neutrophil-lymphocyte ratio, LDH: Lactate dehydrogenase, *: $p < 0.05$

DISCUSSION

NECs are rare, aggressive malignancies with a poor prognosis. The first-line treatment for NECs is platinum plus etoposide therapy.⁵ However, there is no consensus yet on second-line treatment. There is also no randomized study on the second-line treatment of GEP-NECs. In our study, we aimed to investigate the effectiveness of first-line and second-line therapies in patients with GEP-NECs.

When examining similar studies, Yamaguchi et al.'s¹⁶ retrospective study, which included 258 patients diagnosed with gastrointestinal system NECs, aimed to compare the effectiveness of irinotecan plus etoposide and cisplatin plus etoposide regimens, as well as to investigate their effects according to tumor localization. The median OS in all patients was 11.5 months, with median PFS1 ranging from 3.7 to 5.4 months and PFS2 ranging from 1.6 to 2.4 months, depending on the treatment regimen and tumor localization. As a result, it was noted that no clear difference was found between the two treatment regimens; however, larger studies were deemed necessary to determine whether different treatment options could be preferred based on tumor localization in treatment selection. The survival results of this study were similar to the results of our study. Yoon et al.'s¹⁷ study, including 64 patients diagnosed with GEP-NEC, investigated the effectiveness of cisplatin plus etoposide treatment. The PFS was determined to be 3.5 months (range, 3.7-5.3) in the entire patient group. In the study by Heetfeld et al.,¹⁸ including 167 patients with GEP-NECs, the median OS was 16.4 months (95% CI, 13.4-19.5 months), the median PFS1 was 5 months (95% CI, 4.0-6.1 months), and the median PFS2 was 2.98 months (95% CI, 2.56-3.41 months). The survival outcomes of these studies were similar to our study.

In neuroendocrine malignancies, studies have been conducted on factors and various scoring systems that may affect treatment efficacy and prognosis. The most frequently studied markers in this subject are inflammatory markers, including WBC count, neutrophil count, CRP, high-sensitivity CRP, LDH, and aspartate aminotransferase (AST).^{11,19} In addition, the prognostic value of scores such as NLR and PLR, which are derived by combining various laboratory values with these inflammatory markers, has been the subject of numerous studies.^{12,20,21} In these studies, results showed that increased WBC and neutrophil counts, as well as high NLR and PLR scores, are poor prognostic indicators. In our study, we found that increased NLR scores were associated with worse overall survival. The median NLR value in our study was consistent with the values reported in the literature.^{22,23} Upon examining numerous studies and meta-analyses, the cut-off value of PLR was determined to be between 150 and 300,²⁴ and our median PLR value fell within this range. Hypothetically, the ability to distinguish between good and poor prognostic patients with these scores may also affect treatment strategies. Although there are multiple treatment options in second-line treatment, none of them has become standard treatment. Additionally, there is no randomized study directly comparing these treatment approaches with each other in the current literature.

Targeted therapies have begun to play an important role in almost all cancer types with the increase in tumor-agnostic

markers. It is recommended to investigate actionable molecular alterations in the progression of NECs after first-line treatment. The main ones are pembrolizumab²⁵ and dostarlimab²⁶ in tumors with mismatch-repair deficiency (dMMR) or high microsatellite instability (MSI-H), pembrolizumab²⁷ and nivolumab plus ipilimumab²⁸ in tumors with high tumor mutational burden (TMB-H; ≥ 10 mutations per megabase), although data are limited, and the combination of dabrafenib plus trametinib in the presence of BRAF V600E mutation.²⁹ Additionally, less commonly, treatments targeting NTRK and RET fusions may also be considered.³⁰⁻³² However, the available data on these treatment options are limited, and actionable molecular alterations are not detected in the majority. In this patient group, the treatment decision can be challenging due to the lack of a standard treatment and limited treatment options. In the second-line treatment, similar to small-cell lung cancer treatment, platinum plus etoposide rechallenge treatment can be tried in cases of progression that develop more than 6 months after platinum plus etoposide treatment. However, in cases of progression that develop within 6 months, the current NCCN guidelines recommend capecitabine plus temozolomide, platinum-based chemotherapy plus irinotecan, FOLFOX, FOLFIRI, and FOLFIRINOX as treatment options. In a retrospective study conducted in Japan, which included 533 patients diagnosed with metastatic gastroenteropancreatic neuroendocrine neoplasm, it was shown that approximately half of the patients received treatments that were not recommended in the guidelines.³³ This situation has probably developed due to the lack of a standard treatment recommendation.

Ongoing clinical trials are investigating the efficacy of tarlatamab treatment, which has recently begun to gain traction in the treatment of small cell lung cancer, a type of NEC, as well as in the treatment of pulmonary and extrapulmonary neuroendocrine carcinoma.^{34,35} A significant portion of these studies investigates the efficacy of tarlatamab in second-line therapy in patients who have received platinum-based chemotherapy in first-line treatment. The fact that there is no satisfactory second-line treatment may change as a result of these studies. However, to our knowledge, there is no ongoing study specifically in GEP-NEC patients.

Limitations

Our study has several limitations, including the small number of patients and its retrospective, non-randomized nature. However, considering that a significant portion of patients with NEC can not receive second-line treatment due to the aggressive nature of the disease and limited treatment options, the number of patients in our study cannot be considered insignificant. Additionally, the fact that molecular tests, such as MSI or BRAF, were not performed for targeted therapies in any patient during the planning of first-line or second-line treatment may be considered another limitation. Furthermore, there are no randomized studies on this topic in the existing literature.

CONCLUSION

As a result, our study found that higher NLR scores and Ki67 values were significantly associated with worse median

OS. Survival outcomes were similar to those reported in the literature for both OS, PFS1, and PFS2. In the management of NECs, new treatment approaches and prognostic markers are needed due to the lack of standard treatment recommendations, poor prognosis, and short survival times.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Dr. Abdurrahman Yurtaslan Ankara Oncology Training and Research Hospital Non-interventional Clinical Researches Ethics Committee (Date: 19.10.2023, Decision No: 2023-10/97).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Availability of Data and Materials

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

REFERENCES

- Korse CM, Taal BG, van Velthuisen ML, Visser O. Incidence and survival of neuroendocrine tumours in the Netherlands according to histological grade: experience of two decades of cancer registry. *Eur J Cancer*. 2013;49(8):1975-1983. doi:10.1016/j.ejca.2012.12.022
- Boyar Cetinkaya R, Aagnes B, Thiis-Evensen E, Tretli S, Bergsetuen DS, Hansen S. Trends in incidence of neuroendocrine neoplasms in norway: a report of 16,075 cases from 1993 through 2010. *Neuroendocrinology*. 2017;104(1):1-10. doi:10.1159/000442207
- Cho MY, Kim JM, Sohn JH, et al. Current trends of the incidence and pathological diagnosis of gastroenteropancreatic neuroendocrine tumors (GEP-NETs) in Korea 2000-2009: multicenter study. *Cancer Res Treat*. 2012;44(3):157-165. doi:10.4143/crt.2012.44.3.157
- Dasari A, Mehta K, Byers LA, Sorbye H, Yao JC. Comparative study of lung and extrapulmonary poorly differentiated neuroendocrine carcinomas: a SEER database analysis of 162,983 cases. *Cancer*. 2018;124(4):807-815. doi:10.1002/cncr.31124
- Mitry E, Baudin E, Ducreux M, et al. Treatment of poorly differentiated neuroendocrine tumours with etoposide and cisplatin. *Br J Cancer*. 1999;81(8):1351-1355. doi:10.1038/sj.bjc.6690325
- Moertel CG, Kvols LK, O'Connell MJ, Rubin J. Treatment of neuroendocrine carcinomas with combined etoposide and cisplatin. Evidence of major therapeutic activity in the anaplastic variants of these neoplasms. *Cancer*. 1991;68(2):227-232. doi:10.1002/1097-0142(19910715)68:2<227::aid-cncr2820680202>3.0.co;2-i
- Andreatos N, McGarrah PW, Sonbol MB, et al. Managing metastatic extrapulmonary neuroendocrine carcinoma after first-line treatment. *Curr Oncol Rep*. 2023;25(10):1127-1139. doi:10.1007/s11912-023-01438-w
- Sorbye H, Welin S, Langer SW, et al. Predictive and prognostic factors for treatment and survival in 305 patients with advanced gastrointestinal neuroendocrine carcinoma (WHO G3): the NORDIC NEC study. *Ann Oncol*. 2013;24(1):152-160. doi:10.1093/annonc/mds276
- Panzuto F, Merola E, Pavel ME, et al. Stage IV gastro-entero-pancreatic neuroendocrine neoplasms: a risk score to predict clinical outcome. *Oncologist*. 2017;22(4):409-415. doi:10.1634/theoncologist.2016-0351
- Nießen A, Schimmack S, Sandini M, et al. C-reactive protein independently predicts survival in pancreatic neuroendocrine neoplasms. *Scientific Reports*. 2021;11(1):23768. doi:10.1038/s41598-021-03187-x
- Zou J, Li Q, Kou F, et al. Prognostic value of inflammation-based markers in advanced or metastatic neuroendocrine tumours. *Curr Oncol*. 2019;26(1):e30-e38. doi:10.3747/co.26.4135
- Grenader T, Pavel ME, Ruzsiewicz PB, et al. Prognostic value of the neutrophil/lymphocyte ratio in enteropancreatic neuroendocrine tumors. *Anticancer Drugs*. 2020;31(3):216-222. doi:10.1097/cad.0000000000000909
- Shirasawa M, Yoshida T, Horinouchi H, et al. Prognostic impact of peripheral blood neutrophil to lymphocyte ratio in advanced-stage pulmonary large cell neuroendocrine carcinoma and its association with the immune-related tumour microenvironment. *Br J Cancer*. 2021;124(5):925-932. doi:10.1038/s41416-020-01188-7
- Kim HJ, Lee KH, Shim HJ, et al. Prognostic significance of the neutrophil-lymphocyte ratio and platelet-lymphocyte ratio in neuroendocrine carcinoma. *Chonnam Med J*. 2022;58(1):29-36. doi:10.4068/cmj.2022.58.1.29
- Shi M, Zhao W, Zhou F, et al. Neutrophil or platelet-to-lymphocyte ratios in blood are associated with poor prognosis of pulmonary large cell neuroendocrine carcinoma. *Translat Lung Cancer Res*. 2020;9(1):45-54.
- Yamaguchi T, Machida N, Morizane C, et al. Multicenter retrospective analysis of systemic chemotherapy for advanced neuroendocrine carcinoma of the digestive system. *Cancer Sci*. 2014;105(9):1176-1181. doi:10.1111/cas.12473
- Yoon SE, Kim JH, Lee SJ, et al. The impact of primary tumor site on outcomes of treatment with etoposide and cisplatin in grade 3 gastroenteropancreatic neuroendocrine carcinoma. *J Cancer*. 2019;10(14):3140-3144. doi:10.7150/jca.30355
- Heetfeld M, Chougnet CN, Olsen IH, et al. Characteristics and treatment of patients with G3 gastroenteropancreatic neuroendocrine neoplasms. *Endocr Relat Cancer*. 2015;22(4):657-664. doi:10.1530/erc-15-0119
- Freis P, Grailot E, Rousset P, et al. Prognostic factors in neuroendocrine carcinoma: biological markers are more useful than histomorphological markers. *Scientific Reports*. 2017;7(1):40609. doi:10.1038/srep40609
- Matsas S, Junior PNA, Giglio AD. Prognostic role of platelet-to-lymphocyte ratio (PLR) and neutrophil-to-lymphocyte ratio (NLR) in advanced gastric cancer treated with immunotherapy: a systematic review and meta-analysis. *J Clin Oncol*. 2024;42(3_suppl):397-397. doi:10.1200/JCO.2024.42.3_suppl.397
- Liu T, Chen X, Mo S, et al. Molecular subtypes and prognostic factors of lung large cell neuroendocrine carcinoma. *Transl Lung Cancer Res*. 2024;13(9):2222-2235. doi:10.21037/tlcr-24-292
- Panni RZ, Lopez-Aguilar AG, Liu J, et al. Association of preoperative monocyte-to-lymphocyte and neutrophil-to-lymphocyte ratio with recurrence-free and overall survival after resection of pancreatic neuroendocrine tumors (US-NETSG). *J Surg Oncol*. 2019;120(4):632-638. doi:https://doi.org/10.1002/jso.25629
- Abdelmalak R, Lythgoe MP, Evans J, et al. Exploration of novel prognostic markers in grade 3 neuroendocrine neoplasia. *Cancers*. 2021;13(16):4232.
- Zhou X, Du Y, Huang Z, et al. Prognostic value of PLR in various cancers: a meta-analysis. *PLoS One*. 2014;9(6):e101119. doi:10.1371/journal.pone.0101119
- Marabelle A, Le DT, Ascierto PA, et al. Efficacy of pembrolizumab in patients with noncolorectal high microsatellite instability/mismatch repair-deficient cancer: results from the phase II KEYNOTE-158 study. *J Clin Oncol*. 2020;38(1):1-10. doi:10.1200/JCO.19.02105
- Andre T, Berton D, Curigliano G, et al. Antitumor activity and safety of dostarlimab monotherapy in patients with mismatch repair deficient solid tumors: a nonrandomized controlled trial. *JAMA Netw Open*. 2023;6(11):e2341165. doi:10.1001/jamanetworkopen.2023.41165

27. Marabelle A, Fakih M, Lopez J, et al. Association of tumour mutational burden with outcomes in patients with advanced solid tumours treated with pembrolizumab: prospective biomarker analysis of the multicohort, open-label, phase 2 KEYNOTE-158 study. *Lancet Oncol.* 2020;21(10):1353-1365. doi:10.1016/S1470-2045(20)30445-9
28. Mohamed A, Vijayvergia N, Kurian M, Liu L, Fu P, Das S. Exploring real world outcomes with nivolumab plus ipilimumab in patients with metastatic extra-pulmonary neuroendocrine carcinoma (EP-NEC). *Cancers (Basel).* 2022;14(11):2695. doi:10.3390/cancers14112695
29. Salama AKS, Li S, Macrae ER, et al. Dabrafenib and trametinib in patients with tumors with BRAF(V600E) mutations: results of the NCI-MATCH trial subprotocol H. *J Clin Oncol.* 2020;38(33):3895-3904. doi:10.1200/JCO.20.00762
30. Subbiah V, Wolf J, Konda B, et al. Tumour-agnostic efficacy and safety of selpercatinib in patients with RET fusion-positive solid tumours other than lung or thyroid tumours (LIBRETTO-001): a phase 1/2, open-label, basket trial. *Lancet Oncol.* 2022;23(10):1261-1273. doi:10.1016/s1470-2045(22)00541-1
31. Hong DS, DuBois SG, Kummar S, et al. Larotrectinib in patients with TRK fusion-positive solid tumours: a pooled analysis of three phase 1/2 clinical trials. *Lancet Oncol.* 2020;21(4):531-540. doi:10.1016/s1470-2045(19)30856-3
32. Demetri GD, De Braud F, Drilon A, et al. Updated integrated analysis of the efficacy and safety of entrectinib in patients with NTRK fusion-positive solid tumors. *Clin Cancer Res.* 2022;28(7):1302-1312. doi:10.1158/1078-0432.Ccr-21-3597
33. Yamamoto S, Sakakibara N, Hirano H, et al. The real-world selection of first-line systemic therapy regimen for metastatic gastroenteropancreatic neuroendocrine neoplasm in Japan. *Sci Rep.* 2022;12(1):17601. doi:10.1038/s41598-022-22718-8
34. Hospital NTU. Tarlatamab for advanced extrapulmonary small cell carcinoma and neuroendocrine carcinoma (TAURUS). 2025.
35. Inkeun Park AMCRP. Efficacy and safety evaluation of tarlatamab in advanced extrapulmonary neuroendocrine carcinoma patients (DeLLight). 2025.

The impact of perceived social support, internalized stigma, and childhood trauma on quality of life in schizophrenia spectrum disorders: a mediation and moderation approach

 Ulaş Korkmaz,  Fatma Gül Helvacı Çelik

Department of Psychiatry, Faculty of Medicine, Giresun University, Giresun, Türkiye

Cite this article as: Korkmaz U, Helvacı Çelik FG. The impact of perceived social support, internalized stigma, and childhood trauma on quality of life in schizophrenia spectrum disorders: a mediation and moderation approach. *J Health Sci Med.* 2025;8(6):966-974.

Received: 25.07.2025

Accepted: 13.09.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to determine whether perceived social support and internalized stigma mediate the association between symptom severity and quality of life (QoL) in individuals with schizophrenia spectrum disorders (SSDs) and whether childhood trauma moderate this relationship.

Methods: A cross-sectional design was used with 258 patients diagnosed with schizophrenia, schizoaffective disorder, or unspecified psychosis at a community mental health center. Standardized scales were administered to assess QoL, perceived social support, internalized stigma, history of childhood trauma, and symptom severity. Mediation analyses were conducted to test the indirect effects of symptom severity on QoL through social support and internalized stigma. Moderation analyses examined the effect of childhood traumas on the relationship between symptom severity and QoL.

Results: Higher symptom severity was associated with lower perceived social support and greater internalized stigma, which in turn predicted worse QoL. Perceived social support and internalized stigma significantly mediated the negative effect of symptom severity on QoL. Childhood trauma had a significant moderating effect: the negative association between symptom severity and QoL was stronger among patients with a history of childhood trauma. These effects remained significant even after controlling for other variables.

Conclusion: Psychosocial factors significantly influence QoL in patients with SSDs. Perceived social support and self-stigma emerged as key mediators linking symptom severity to QoL, and a history of childhood trauma exacerbated the impact of symptoms on QoL. Interventions that strengthen social support, reduce internalized stigma and address childhood traumatic experiences may be crucial for improving QoL in the SSD population.

Keywords: Schizophrenia spectrum disorders, quality of life, perceived social support, internalized stigma, childhood trauma

INTRODUCTION

Quality of life (QoL) is closely associated with individuals' subjective well-being, happiness, and sense of purpose and is considered a fundamental component of both health and human existence. Standard indicators of QoL include financial circumstances, employment status, environmental conditions, physical and mental health, access to education, opportunities for recreation and leisure, social connectedness, spiritual beliefs, personal safety, and freedom.^{1,2} Higher levels of QoL and a sense of meaning have been shown to increase individuals' psychological well-being, help them cope with stress, and even extend their lifespan.³

Mental disorders negatively impact individuals' success in areas such as education, employment, and social participation, and people suffering from these disorders often report a lower QoL than the general population.^{4,5} Severe mental disorders such as schizophrenia spectrum disorders (SSDs; including schizophrenia, schizoaffective disorder, or unspecified

psychosis) tend to have a chronic course and reduce QoL of individuals by seriously affecting their thoughts, emotions, behaviors, and social and occupational functionality.^{6,7} In this context, controlling SSD symptoms is considered crucial for recovery and improving QoL in individuals with SSDs.

When considering psychosocial factors that may affect QoL, perceived social support is one of the most prominent protective factors. Perceived social support refers to individuals' belief that they can receive emotional, informational, and instrumental help from people around them. This perception strengthens individuals' stress-coping mechanisms and increases their overall well-being.⁸⁻¹⁰ In particular, SSDs may reduce perceived social support by causing deterioration in social relationships.^{8,11-13} Social support is considered a protective factor that reduces the severity of psychiatric symptoms.¹⁴ This is confirmed by the fact that patients with high levels of social support experience fewer symptoms and

Corresponding Author: Ulaş Korkmaz, ulas.korkmaz@giresun.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

significantly higher overall QoL scores.¹⁵ Therefore, it can be considered that perceived social support may play a mediating role in the effect of symptom severity on QoL.

Internalized stigma is another important factor that negatively affects patients' well-being. Internalized stigma in mental disorders is the practice of individuals applying society's negative beliefs and attitudes towards mental disorders to themselves, which reduces self-esteem, self-efficacy, and hope.^{11,16} This process can lead individuals to develop a negative self-identity, withdraw from society, and consequently reduce their QoL.¹⁵ SSDs increase the risk of internalized stigma, mainly due to severe symptoms and chronic course.^{11,17,18} Therefore, it can be argued that the level of stigma may also play a mediating role in the relationship between symptoms and QoL.

Childhood trauma refers to negative and distressing experiences in early life, such as physical, emotional, or sexual abuse, neglect, domestic violence, parental loss, or serious illness.¹⁹ Research shows that childhood trauma is an important risk factors for the development of mental disorders, especially SSDs.²⁰ In individuals with a history of childhood trauma, SSDs tend to develop earlier, present with more severe symptoms, and show greater resistance to treatment.^{21,22} QoL levels of patients with a history of trauma were found to be significantly lower than those who had not experienced trauma.²³ These findings highlight the lasting effects of childhood trauma on mental health. Moreover, it is conceivable that childhood trauma may moderate the course of SSDs.

This study aims to examine the mediating roles of perceived social support and internalized stigma in the relationship between symptom severity and QoL among individuals with SSDs, as well as the moderating effect of childhood trauma within this relationship, by addressing these variables within a comprehensive model. In this way, the study seeks to contribute to the literature by revealing the complex mechanisms underlying the relationship between SSD symptom severity and QoL and to provide a foundation for the development of more targeted, holistic, and effective intervention strategies for this population.

The following hypotheses were tested in this study:

- **Hypothesis 1:** Perceived social support has a mediating effect on the relationship between SSD symptom severity and QoL.
- **Hypothesis 2:** Internalized stigma has a mediating effect on the relationship between SSD symptom severity and QoL.
- **Hypothesis 3:** As the level of childhood trauma increases, the negative relationship between SSD symptom severity and QoL will strengthen.

METHODS

Study Design and Data Collection

Ethical approval was received for this research from the the Giresun Training and Research Hospital Clinical Researches Ethics Committee (Date: 02.05.2025, Decision No:

30.04.2025/02). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

This study was conducted in a community mental health center. The study population consisted of patients who presented to the community mental health center and were diagnosed with schizophrenia, schizoaffective disorder, or unspecified psychosis according to DSM-5-TR. Patients whose participation was approved both by themselves and their legal guardians and who signed the informed consent form were included in the study. The study was conducted with 258 participants. The inclusion criteria were determined to be over 18 years of age, having the mental capacity to understand and respond to instructions, being willing to participate in the study, and obtaining written consent from both the participants and their legal guardians. Exclusion criteria were determined as having any psychiatric diagnosis other than the diagnoses specified, not giving consent to participate in the study, being in an acute psychotic state, showing symptoms such as agitation, aggression, or uncooperativeness at a level that would prevent participation in the study, having a loss of cognitive function at a level that would prevent participation in the study, having a neurological disease that affects the central nervous system and significantly affects cognitive or behavioral functions (e.g., epilepsy, traumatic brain injury), and having a severe and uncontrolled physical disease that would prevent participation in the study or pose a risk (e.g., terminal cancer, decompensated heart failure).

A sociodemographic data form, the World Health Organization Quality of Life Instrument-Short Form (WHOQOL-BREF), Multidimensional Scale of Perceived Social Support (MSPSS), Internalized Stigma of Mental Illness Scale (ISMI), Childhood Trauma Questionnaire (CTQ) and Positive and Negative Syndrome Scale (PANSS) were applied to each participant. Written and verbal consent was obtained from the patients and their guardians prior to the administration of the scales.

Measures

A structured sociodemographic data form developed by the researchers was used to assess the demographic and clinical characteristics of the participants. Information was collected regarding the participants' age, gender, marital status, duration of education, employment status, place of residence, monthly income (calculated as the household's total monthly income divided by the number of individuals in the household, and then expressed as a ratio to the national minimum wage.), duration of the disorder, and the presence of comorbid physical illnesses. In addition, information on participants' current antipsychotic medication use was also assessed. In this context, variables such as clozapine use, long-acting injectable (LAI) antipsychotic treatment, and antipsychotic treatment regimen (monotherapy/polypharmacy) were included in the form.

The World Health Organization Quality of Life Instrument-Short Form is a scale designed to assess individuals' quality of life across four dimensions: physical health, psychological health, social relationships, and the environment.² WHOQOL-BREF scoring is done with Likert-type responses ranging

from 1 to 5, and each dimension score is calculated. While the physical health dimension encompasses energy level, pain, sleep, and daily activities, the psychological health dimension reflects mood, self-perception, mindset, and mental well-being. The social relationships dimension assesses individuals' perception of social support and satisfaction with their close relationships. The environment dimension encompasses environmental factors such as individuals' living conditions, access to health services, security, and leisure opportunities. The Turkish validity and reliability study of WHOQOL-BREF has been conducted.²⁴

Multidimensional Scale of Perceived Social Support was used to measure the perceived social support level of individuals. The scale, developed by Zimet et al.,²⁵ consists of 12 items. Each item is answered on a 7-Point Likert Scale, with higher scores indicating greater perceived support. The Turkish adaptation of the scale was carried out by Eker and Arkar.²⁶

Internalized Stigma of Mental Illness Scale was used to assess the patients' level of internalized stigma. The ISMI, developed by Ritsher et al.,²⁷ consists of 29 statements. Individuals respond to each statement on a scale of 1 to 4. High scores indicate high levels of self-stigma due to mental disorders. The adaptation study of the scale in Türkiye was conducted by Ersoy and Varan.²⁸

Childhood Trauma Questionnaire was used to screen participants' childhood traumatic experiences. CTQ, developed by Bernstein et al.,¹⁹ consists of a total of 28 questions in five sub-dimensions: emotional, physical, sexual abuse, and emotional and physical neglect. Emotional abuse includes verbal violence, such as humiliation and threats; physical abuse includes bodily harm; and sexual abuse includes sexual behaviors directed at children. Emotional neglect includes the child experiencing a lack of attention and love, and physical neglect includes the failure to meet basic life needs. Each item is answered in a 5-point Likert format. Higher scores indicate higher levels of exposure to the relevant type of trauma. The Turkish validity and reliability study was conducted by Şar et al.²⁹

Positive and Negative Syndrome Scale is a structured clinical interview and observation scale widely used to assess the severity of SSD symptoms.³⁰ The scale evaluates three main areas: positive symptoms (hallucinations, delusions), negative symptoms (affective blunting, social withdrawal), and general psychopathology. The scale is administered and scored by a trained clinician. Each item is scored between 1 and 7, with higher total scores indicating more severe symptoms. The validity and reliability study of the scale in Turkish has been conducted.³¹

Statistical Analysis

IBM SPSS Statistics 27.0 and PROCESS Macro were used in statistical analyses. Descriptive values are presented as frequency-percentage distributions and mean-standard deviation values. Within the scope of the primary analyses, mediation and moderation analyses were conducted. The significance of the indirect effect was assessed with bootstrap confidence intervals (5,000 resamples). In all mediation and moderation analyses, gender, age, marital status, duration of education, employment status, place of residence, duration of

the disorder, LAI use, clozapine use, antipsychotic treatment regimen, presence of comorbid physical illnesses, and monthly income level were added as control variables. Analysis outputs were reported, including the indirect effect of symptom severity on QoL, effect sizes (K^2) of mediator variables, and explained variances (R^2) of the models. The results were evaluated at a 95% confidence level, and $p < 0.05$ was taken as the criterion for statistical significance.

RESULTS

Distribution of Categorical Variables in the Sample

The sample consisted of 61.6% males. Among the participants, 78.7% were single, 85.7% were unemployed, 74% resided in urban areas, 34.5% had a comorbid physical illness, and 41.1% were receiving LAI antipsychotic treatment. 16.3% of the participants were using clozapine. 59.3% of the patients were using more than one antipsychotic drug, and 40.7% were receiving a single antipsychotic treatment (Table 1).

		n=258	%
Gender	Female	99	38.4
	Male	159	61.6
Marital status	Married	55	21.3
	Single	203	78.7
Employment status	Employed	37	14.3
	Unemployed	221	85.7
Place of residence	Urban	191	74
	Rural	67	26
Comorbid physical illness	No	169	65.5
	Yes	89	34.5
LAI use	No	152	58.9
	Yes	106	41.1
Clozapine use	No	216	83.7
	Yes	42	16.3
Antipsychotic treatment regimen	Monotherapy	105	40.7
	Polypharmacy	153	59.3
LAI: Long-acting injectable			

Distribution of Numerical Variables in the Sample

The mean age of the patients was 46.67 years (SD=12.76), and the mean duration of education was 8.79 years (SD=4.34). The mean duration of the disorder was 19.31 years (SD=11.08), and the monthly income was 0.74 times the minimum wage (SD=0.43). The mean WHOQOL-BREF total score was 86.95 (SD=15.95). The mean MSPSS score was 43.60 (SD=18.14), and the mean ISMI score was 63.59 (SD=14.87). The mean CTQ total score was 44.24 (SD=14.78) (Table 2).

The Mediating Effect of Perceived Social Support in the Relationship between Symptom Severity and QoL

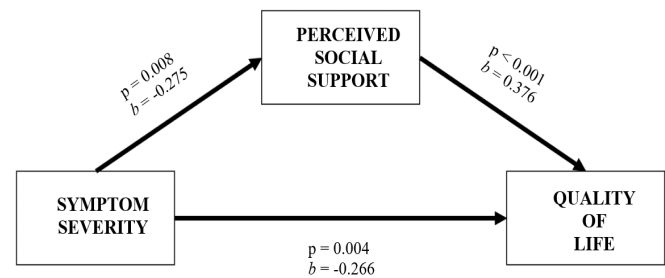
Perceived social support was found to have a significant mediating role in the relationship between SSD symptom severity and QoL (Figure 1). Analysis results showed that symptom severity had a significant effect on perceived social support ($b = -0.275$, $p = 0.008$). The perceived social support score significantly predicted the total QoL score ($b = 0.376$,

Table 2. Numerical variables in the sample

	Mean	SD
Age	46.67	12.76
Duration of education (years)	8.79	4.34
Monthly income (ratio to minimum wage)	0.74	0.43
Duration of the disorder (years)	19.31	11.08
WHOQOL-BREF total	86.95	15.95
General health	6.34	1.74
Physical health	23.90	4.93
Psychological health	19.62	4.29
Social relationships	8.21	2.65
Environment	28.88	5.77
MSPSS	43.60	18.14
ISMI	63.59	14.87
CTQ total	44.24	14.78
Emotional abuse	8.67	3.67
Physical abuse	7.00	3.53
Physical neglect	9.33	3.37
Emotional neglect	12.66	4.71
Sexual abuse	6.58	3.33
PANSS total	50.48	12.26
Positive symptoms	10.91	3.06
Negative symptoms	13.28	4.02
General psychopathology	26.29	6.72

SD: Standard deviation, WHOQOL-BREF: The World Health Organization Quality of Life Instrument-Short Form, MSPSS: Multidimensional Scale of Perceived Social Support, ISMI: Internalized Stigma of Mental Illness Scale, CTQ: Childhood Trauma Questionnaire, PANSS: Positive and Negative Syndrome Scale

$p < 0.001$). The initially significant direct effect of symptom severity on QoL was reduced to non-significance when perceived social support was included as a mediator ($b = -0.163$, 95% CI = $-0.326 - 0.0002$). In contrast, the indirect effect of symptom severity on QoL was found to be significant through perceived social support ($b = -0.103$, 95% CI = $-0.187 - -0.030$). This finding suggests that the relationship between symptom severity and QoL is explained by perceived social support. The mediation effect size (K^2) was 0.079, and the model explained approximately 25% of the variance in the total QoL score ($R^2 = 0.253$). The mediation effect remained significant when ISMI and CTQ variables were added to the model along with other control variables ($b = -0.046$, 95% CI = $-0.097 - -0.005$).



Direct effect: $b = -0.163$, 95% CI = $-0.326 - 0.0002$

Indirect effect: $b = -0.103$, 95% CI = $-0.187 - -0.030$

$K^2 = 0.079$, $R^2 = 0.253$

b : Unstandardized beta coefficient, p : Significance level

Figure 1. Mediating effect of perceived social support on the relationship between symptom severity and QoL

QoL: Quality of life

For each sub-dimension of QoL, the mediation relationships between PANSS positive, negative, and general psychopathology symptom clusters and perceived social support were also examined (Table 3). The effects of positive and negative symptoms on all QoL sub-dimensions were mediated by perceived social support. In contrast, no significant indirect effect of general psychopathology symptoms through perceived social support was found for any QoL sub-dimension.

The Mediating Effect of Internalized Stigma on the Relationship between Symptom Severity and QoL

Internalized stigma was also found to be a significant mediator in the effect of symptom severity on QoL (Figure 2). The PANSS total score significantly predicted the level of internalized stigma ($b = 0.326$, $p < 0.001$). The level of internalized stigma was found to have a significant effect on the QoL score ($b = -0.565$, $p < 0.001$). When the internalized stigma mediator was included in the model, the direct effect of symptom severity on QoL ceased to be significant ($b = -0.082$, 95% CI = $-0.240 - 0.077$). The indirect effect of symptom severity on QoL was found to be significant through internalized stigma ($b = -0.184$, 95% CI = $-0.285 - -0.097$). This result indicates that the negative effect of symptom severity on QoL is largely mediated by internalized stigma. K^2 was 0.141, and the model explained 31.8% of the variance in QoL overall ($R^2 = 0.318$). The mediation effect remained significant when MSPSS and CTQ were added to the model as control variables ($b = -0.124$, 95% CI = $-0.196 - -0.058$).

Table 3. Mediating effect of perceived social support on the effect of symptom clusters on QoL sub-dimensions

X	Y	b	95% CI	K^2	R^2
Positive symptoms	Physical health	-0.088	-0.712 - -0.018	0.054	0.220
Negative symptoms	Physical health	-0.117	-0.194 - -0.057	0.096	0.211
Positive symptoms	Psychological health	-0.085	-0.165 - -0.017	0.061	0.225
Negative symptoms	Psychological health	-0.111	-0.180 - -0.053	0.104	0.217
Positive symptoms	Social relationships	-0.067	-0.134 - -0.012	0.078	0.330
Negative symptoms	Social relationships	-0.087	-0.140 - -0.045	0.132	0.327
Positive symptoms	Environment	-0.095	-0.186 - -0.020	0.051	0.166
Negative symptoms	Environment	-0.128	-0.215 - -0.060	0.089	0.159

QoL: Quality of life, X: Independent variable, Y: Outcome variable, b: Unstandardized beta coefficient, CI: Confidence interval, K^2 : Effect size, R^2 : Coefficient of determination

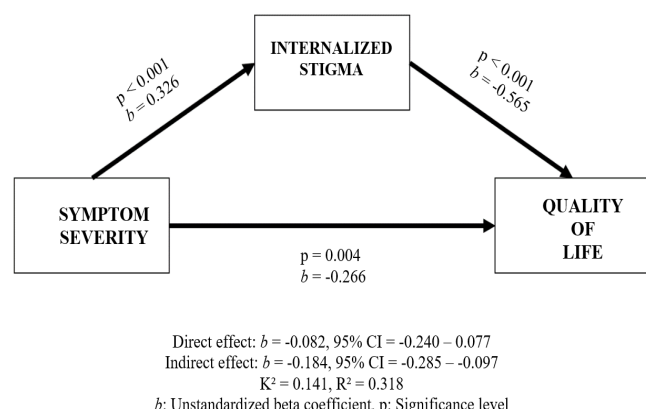


Figure 2. Mediating effect of internalized stigma on the relationship between symptom severity and QoL

QoL: Quality of life

For each sub-dimension of QoL, the mediation relationships between PANSS positive, negative, and general psychopathology symptom clusters and internalized stigma were also examined (Table 4). In all QoL sub-dimensions, the effects of positive and negative symptoms and general psychopathology were mediated by internalized stigma.

The Moderating Effect of Childhood Trauma on the Relationship between Symptom Severity and QoL

Childhood trauma was found to have a significant moderating effect on the relationship between SSD symptom severity and QoL. The interaction effect (PANSS×CTQ) was statistically significant ($b = -0.015$, 95%CI= -0.027 - -0.003). This result suggests that the negative relationship between symptom severity and QoL is more pronounced among individuals with higher CTQ scores. In other words, as childhood trauma increases, the negative impact of symptom severity on QoL becomes stronger. Figure 3 illustrates the moderating effect. In patients with high levels of childhood trauma, increases in PANSS scores were associated with a steeper decline in QoL compared to those with low levels of trauma. When ISMI and MSPSS were included in the model alongside other control variables, the moderating effect remained significant ($b = -0.015$, 95%CI= -0.025 - -0.005).

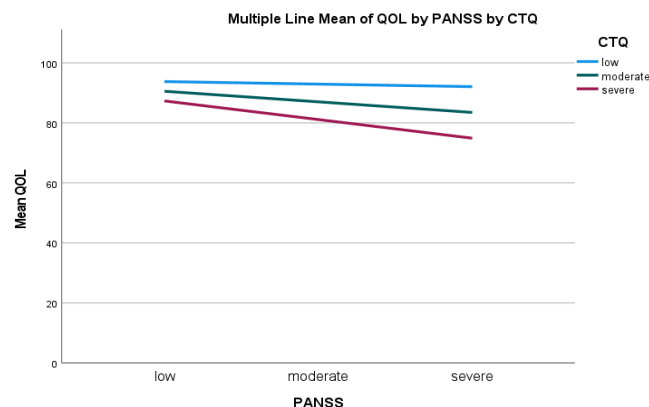


Figure 3. The moderating effect of childhood trauma on the relationship between symptom severity and QoL

QoL: Quality of life, PANSS: Positive and Negative Syndrome Scale, CTQ: Childhood Trauma Questionnaire

For each sub-dimension of QoL, the moderating effects of PANSS positive, negative, and general psychopathology symptom clusters and childhood traumas were also examined (Table 5). A history of childhood trauma strengthened the associations between specific symptom clusters and QoL. The effects of positive symptoms and general psychopathology on the physical and psychological health dimensions were more negative among individuals with a history of emotional, physical, or sexual abuse. No significant moderating effect of childhood trauma on the physical health and psychological health was found for negative symptoms. Relationships between positive and negative symptoms and the social relationships dimension were not affected any trauma types. However, the relationship between general psychopathology level and social relationships was more negative in individuals with a history of emotional abuse and physical abuse. The effects of positive symptoms and general psychopathology on the environment dimension of QoL were more negative among individuals with a history of emotional, physical, or sexual abuse. In addition, the effect of negative symptoms on the environment dimension of QoL was found to be stronger among individuals with a history of emotional or sexual abuse. However, emotional and physical neglect scores did not significantly moderate the relationship between any symptom group and QoL.

Table 4. Mediating effect of internalized stigma on the effect of symptom clusters on QoL sub-dimensions

X	Y	b	95% CI	K ²	R ²
Positive symptoms	Physical health	-0.171	-0.274 - -0.085	0.106	0.264
Negative symptoms	Physical health	-0.114	-0.200 - -0.042	0.093	0.066
General psychopathology	Physical health	-0.074	-0.121 - -0.034	0.102	0.260
Positive symptoms	Psychological health	-0.192	-0.300 - -0.096	0.137	0.341
Negative symptoms	Psychological health	-0.126	-0.216 - -0.047	0.118	0.345
General psychopathology	Psychological health	-0.082	-0.129 - -0.040	0.129	0.337
Positive symptoms	Social relationships	-0.076	-0.131 - -0.032	0.088	0.221
Negative symptoms	Social relationships	-0.049	-0.092 - -0.015	0.075	0.235
General psychopathology	Social relationships	-0.033	-0.057 - -0.014	0.083	0.216
Positive symptoms	Environment	-0.184	-0.305 - -0.085	0.097	0.201
Negative symptoms	Environment	-0.122	-0.217 - -0.046	0.085	0.201
General psychopathology	Environment	-0.079	-0.131 - -0.038	0.093	0.066

QoL: Quality of life, X: Independent variable, Y: Outcome variable, b: Unstandardized beta coefficient, CI: Confidence interval, K²: Effect size, R²: Coefficient of determination

Table 5. Moderating effects of childhood traumas on the effects of symptom clusters on QoL sub-dimensions

X → Y	W	b	95% CI
Positive symptoms → Physical health	Emotional abuse	-0.072	-0.130 - -0.013
	Physical abuse	-0.073	-0.139 - -0.007
	Sexual abuse	-0.092	-0.165 - -0.019
General psychopathology → Physical health	Emotional abuse	-0.050	-0.077 - -0.023
	Physical abuse	-0.055	-0.090 - -0.019
	Sexual abuse	-0.065	-0.103 - -0.027
Positive symptoms → Psychological health	Emotional abuse	-0.073	-0.124 - -0.023
	Physical abuse	-0.082	-0.140 - -0.024
	Sexual abuse	-0.092	-0.156 - -0.027
General psychopathology → Psychological health	Emotional abuse	-0.045	-0.069 - -0.021
	Physical abuse	-0.051	-0.082 - -0.019
	Sexual abuse	-0.055	-0.089 - -0.021
General psychopathology → Social relationships	Emotional abuse	-0.017	-0.032 - -0.001
	Physical abuse	-0.021	-0.041 - -0.002
	Sexual abuse	-0.021	-0.041 - -0.002
Positive symptoms → Environment	Emotional abuse	-0.135	-0.201 - -0.068
	Physical abuse	-0.111	-0.189 - -0.034
	Sexual abuse	-0.148	-0.233 - -0.062
Negative symptoms → Environment	Emotional abuse	-0.056	-0.104 - -0.009
	Sexual abuse	-0.079	-0.141 - -0.016
	Physical abuse	-0.065	-0.096 - -0.034
General psychopathology → Environment	Emotional abuse	-0.074	-0.116 - -0.033
	Physical abuse	-0.074	-0.116 - -0.033
	Sexual abuse	-0.059	-0.104 - -0.014

QoL: Quality of life, X: Independent variable, Y: Outcome variable, W: Moderator variable, b: Unstandardized beta coefficient of interaction effect, CI: Confidence interval

DISCUSSION

This study examined the mediating effects of perceived social support and internalized stigma and the moderating effect of childhood trauma on the relationship between symptom severity and QoL in individuals with SSDs. The findings show that the effects of positive and negative symptoms on all QoL sub-dimensions are explained through social support, and the effects of positive and negative symptoms and general psychopathology are explained through internalized stigma. Additionally, some types of childhood traumas were found to have a significant moderating effect on the relationship between symptom severity and QoL.

The results of the study indicate that QoL decreases significantly as the severity of symptoms increases in individuals with SSDs. This finding is well documented in the literature. Galuppi et al.³² reported that the worsening of symptoms is one of the main factors that negatively affect QoL in patients with schizophrenia. QoL of patients whose symptoms cannot be managed effectively can be significantly reduced. However, the effect of symptom severity on QoL is not absolute and can be influenced by various factors.

This study demonstrated that the relationship between symptom severity and QoL was mediated by perceived social support. The fact that individuals with severe symptoms perceive less social support may be an important

intermediary mechanism contributing to their reduced QoL. The buffering effect of social support on mental health has long been emphasized in the literature. Social support can improve functioning and QoL, even in people with serious mental disorders, by increasing individuals' capacity to cope with stress and reducing their social isolation.³³ It has been reported that patients with schizophrenia who have strong family and friend support have higher rates of remission and better QoL.³⁴ Our findings, consistent with these data, suggest that inadequate social support exacerbates the negative impact of symptoms on QoL. Particularly in patients with pronounced positive and negative symptoms, the weakening or disruption of social relationships due to these symptoms can have a negative impact on QoL. Perceived social support can intervene at this point, reducing the stress caused by symptoms and increasing the individual's satisfaction with life. Although the relationship between psychotic symptoms and perceived social support, as well as between perceived social support and QoL, is well established,³⁵ this study is one of the few studies to demonstrate the mediating role of perceived social support in these relationships. The findings suggest that social support is as important as symptom management in treating SSDs and that psychosocial interventions should focus on strengthening this support.

Another important result of the study is that internalized stigma plays a significant mediating role in the relationship between symptom severity and QoL. Severe symptoms can reinforce feelings of stigma by making individuals feel worthless or socially excluded, which can negatively impact their QoL. This finding is consistent with the literature on the negative consequences of stigma in SSDs. Research shows that a significant proportion of individuals with severe mental disorders experience high levels of internalized stigma, which is a major barrier to recovery.³⁶ Internalized stigma harms both objective and subjective well-being by leading to outcomes such as decreased self-confidence, hopelessness, social withdrawal, and low compliance with treatment.³⁷ There is a reciprocal relationship between psychotic symptoms and stigma. There is a negative relationship between the severity of stigma and QoL.³⁸ Our findings support this model by demonstrating that the effect of symptom severity on QoL is largely mediated by patients' self-stigma. Moreover, to our knowledge, this study is one of the first to investigate the mediating effect of stigma in a single study population, and, in this respect, it makes an important contribution to the literature. This also suggests that combating stigma and interventions aimed at changing patients' negative beliefs about themselves may be at least as important as treating symptoms.

In the study, analyses conducted at the level of sub-dimensions of symptoms showed that the effect of mediating variables was not the same for each symptom type. The mediating effect of perceived social support was found to be statistically significant only for positive and negative symptoms, and no mediation effect of perceived social support was observed for general psychopathology symptoms. In contrast, the mediating role of internalized stigma was significant in all positive, negative, and general symptom categories. This differentiation can be interpreted in light of the literature. As positive symptoms

become more severe, it may become more difficult for patients to maintain social relationships.³⁹ Negative symptoms can directly cause patients to become disconnected from their environment and lose their support resources.⁴⁰ Therefore, the reduction in QoL caused by positive and negative symptoms may be primarily related to patients' lack of social support. General psychopathology symptoms may be one of the factors that reduce QoL, independent of social support. Comorbid symptoms, particularly depression and anxiety,⁴¹ can reduce life satisfaction even if a person has a strong social network. However, it is noteworthy that internalized stigma affects QoL in all symptom subtypes. The feeling of stigma is a common issue that can arise in any patient, regardless of whether positive symptoms, negative symptoms, or general psychopathology predominate.⁴² It has also been reported that depression and anxiety levels increase in patients with high levels of stigma, and feelings of hopelessness and helplessness become more common.^{43,44} Considering these findings, it appears that stigma is a key mechanism that can reduce patient QoL in all symptom patterns. In conclusion, sub-dimension analyses reveal that social support may be a protective factor that plays a role in the effect of positive and negative symptoms on QoL, while stigma may be a general risk factor for all symptom types.

The study found that the negative relationship between symptom severity and QoL was even stronger in SSD patients with a history of childhood trauma. Especially in patients who have experienced emotional, physical, and sexual abuse, the decrease in QoL became more pronounced when symptoms increased compared to those without a history of abuse. These findings are consistent with the literature that a history of trauma worsens the course of illness in patients with schizophrenia. Schizophrenia patients exposed to early trauma have been found to respond less well to treatment, show slower improvement in positive symptoms, and have a lower long-term QoL.⁴⁵ Trauma history affects not only QoL life but also the severity of symptoms. Studies have shown that positive symptoms and general psychopathology symptoms are more severe, and general functionality is more impaired in schizophrenia patients who have experienced abuse in childhood.⁴⁶ Therefore, childhood trauma is an important risk factor that negatively affects the course of SSDs. The unique contribution of this study is that it reveals the moderating role of childhood trauma, that is, a factor that exacerbates the effect of symptoms on QoL. While abuse-type traumas, in particular, showed this effect, a similar moderation effect was not observed for neglect-type traumas. Abuse-type traumas have recently been shown to be more significantly and intensely associated with schizophrenia than neglect-type traumas, and this may be related to the fact that abuse-type traumas cause more brain abnormalities.^{47,48} In this context, the fact that only a history of abuse had a significant effect on the relationship between symptom severity and QoL in our study is consistent with the idea that abuse may pose a more specific risk in SSDs. In conclusion, a history of childhood trauma in SSD patients has a detrimental effect on QoL by both increasing the severity of symptoms and magnifying the harm these symptoms cause to individuals.

The findings suggest that when working with individuals diagnosed with SSDs, it is not enough to focus only on relieving symptoms; psychosocial factors should also be addressed with a holistic approach. In this regard, some implications can be drawn for clinical practice and mental health policies. Patients' social support networks should be strengthened to alleviate the loss of function and QoL deterioration caused by symptoms. Anti-stigma interventions should be a priority at both the clinical and community levels. Because a history of childhood trauma is associated with a poor prognosis in patients with SSD, childhood trauma should be evaluated and treated when necessary. When a history of trauma is detected, trauma-focused psychotherapies should be considered in addition to standard SSD treatment. Recent studies have shown that therapies such as prolonged exposure and eye movement desensitization and reprocessing hold promise for patients with psychotic disorders and trauma-related symptoms.⁴⁹

Limitations

The most striking aspect of this study is that it examines the psychosocial factors affecting QoL of individuals diagnosed with SSDs in a holistic and model-based manner. In particular, addressing the mediating role of perceived social support and internalized stigma in the relationship between symptom severity and QoL and the moderating effect of childhood trauma provides the study with a structure rarely found in the literature. Controlling sociodemographic and clinical variables in the models increased the validity of the findings. In addition, both QoL and symptomatology were evaluated by dividing them into sub-dimensions. These detailed analyses strengthened the clinical significance of the findings by revealing the effects of different symptom groups on different QoL dimensions. Although this study offers original contributions, it has some limitations, such as having a cross-sectional design, collecting data using self-report scales, and a limited sample size.

CONCLUSION

The findings of this study indicate that QoL in individuals with SSDs is significantly affected by psychosocial factors such as internalized stigma and perceived social support. Additionally, childhood trauma was found to have a significant moderating effect on the relationship between symptom severity and QoL. Interventions aimed at reducing internalized stigma and enhancing perceived social support, along with trauma-focused approaches, may play a key role in improving QoL among individuals with SSDs.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Giresun Training and Research Hospital Clinical Researches Ethics Committee (Date: 02.05.2025, Decision No: 30.04.2025/02).

Informed Consent

All patients signed and free and informed consent form.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Boylu AA, Paçacıoğlu B. Quality of life and indicators. *J Acad Res Stud*. 2016;8(15):137-150. doi:10.20990/kilisiibfakademik.266011
- World Health Organization. WHOQOL-BREF: introduction, administration, scoring and generic version of the assessment: field trial version. December 1996. 1996;
- Pergolotti M, Deal AM, Williams GR, et al. Activities, function, and health-related quality of life (HRQOL) of older adults with cancer. *J Geriatr Oncol*. 2017;8(4):249-254. doi:10.1016/j.jgo.2017.02.009
- Defar S, Abraham Y, Reta Y, et al. Health related quality of life among people with mental illness: The role of socio-clinical characteristics and level of functional disability. *Front Public Health*. 2023;11:1134032. doi:10.3389/fpubh.2023.1134032
- Boyer L, Millier A, Perthame E, Aballea S, Auquier P, Toumi M. Quality of life is predictive of relapse in schizophrenia. *BMC Psychiatry*. 2013;13:1-8. doi:10.1186/1471-244X-13-15
- Collaborators GMD. Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990-2019: a systematic analysis for the global burden of disease study 2019. *Lancet Psychiatry*. 2022;9(2):137-150. doi:10.1016/S2215-0366(21)00395-3
- Patmisari E, Huang Y, Orr M, Govindasamy S, Hielscher E, McLaren H. Supported employment interventions with people who have severe mental illness: systematic mixed-methods umbrella review. *PLoS One*. 2024;19(6):e0304527. doi:10.1371/journal.pone.0304527
- Wang J, Mann F, Lloyd-Evans B, Ma R, Johnson S. Associations between loneliness and perceived social support and outcomes of mental health problems: a systematic review. *BMC Psychiatry*. 2018;18:1-16. doi:10.1186/s12888-018-1736-5
- Harandi TF, Taghinasab MM, Nayeri TD. The correlation of social support with mental health: a meta-analysis. *Electronic Physician*. 2017;9(9):5212. doi:10.19082/5212
- Feeney BC, Collins NL. A new look at social support: a theoretical perspective on thriving through relationships. *Person Soc Psychol Rev*. 2015;19(2):113-147. doi:10.1177/1088868314544222
- Livingston JD, Boyd JE. Correlates and consequences of internalized stigma for people living with mental illness: a systematic review and meta-analysis. *Soc Sci Med*. 2010;71(12):2150-2161. doi:10.1016/j.socscimed.2010.09.030
- Thornicroft G, Mehta N, Clement S, et al. Evidence for effective interventions to reduce mental-health-related stigma and discrimination. *Lancet*. 2016;387(10023):1123-1132. doi:10.1016/S0140-6736(15)00298-6
- Eglt GM, Palmer BW, Martin AvS, Tu X, Jeste DV. Loneliness in schizophrenia: Construct clarification, measurement, and clinical relevance. *PLoS One*. 2018;13(3):e0194021. doi:10.1371/journal.pone.0194021
- Lambert M, Naber D. Current issues in schizophrenia: overview of patient acceptability, functioning capacity and quality of life. *CNS drugs*. 2004;18(Suppl 2):5-17. doi:10.2165/00023210-200418002-00002
- Gagiu C, Dionisie V, Manea MC, Mazilu DC, Manea M. Internalised stigma, self-esteem and perceived social support as psychosocial predictors of quality of life in adult patients with schizophrenia. *J Clin Med*. 2024;13(22):6959. doi:10.3390/jcm13226959
- Corrigan PW, Rafacz J, Rüsch N. Examining a progressive model of self-stigma and its impact on people with serious mental illness. *Psychiatr Res*. 2011;189(3):339-343. doi:10.1016/j.psychres.2011.05.024
- Rüsch N, Angermeyer MC, Corrigan PW. Mental illness stigma: concepts, consequences, and initiatives to reduce stigma. *Eur Psychiatr*. 2005;20(8):529-539. doi:10.1016/j.eurpsy.2005.04.004
- Brohan E, Elgie R, Sartorius N, Thornicroft G, Group G-ES. Self-stigma, empowerment and perceived discrimination among people with schizophrenia in 14 European countries: the GAMIAN-Europe study. *Schizophr Res*. 2010;122(1-3):232-238. doi:10.1016/j.schres.2010.02.1065
- Bernstein DP, Fink L, Handelsman L, Foote J. Childhood trauma questionnaire. Assessment of family violence: a handbook for researchers and practitioners. 1998.
- Varese F, Smeets F, Drukker M, et al. Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective-and cross-sectional cohort studies. *Schizophr Bull*. 2012;38(4):661-671. doi:10.1093/schbul/sbs050
- Bailey T, Alvarez-Jimenez M, Garcia-Sanchez AM, Hulbert C, Barlow E, Bendall S. Childhood trauma is associated with severity of hallucinations and delusions in psychotic disorders: a systematic review and meta-analysis. *Schizophr Bull*. 2018;44(5):1111-1122. doi:10.1093/schbul/sbx161
- Wells R, Jacomb I, Swaminathan V, et al. The impact of childhood adversity on cognitive development in schizophrenia. *Schizophr Bull*. 2020;46(1):140-153. doi:10.1093/schbul/sbz033
- Andrianarisoa M, Boyer L, Godin O, et al. Childhood trauma, depression and negative symptoms are independently associated with impaired quality of life in schizophrenia. Results from the national FACE-SZ cohort. *Schizophr Res*. 2017;185:173-181. doi:10.1016/j.schres.2016.12.021
- Eser E, Fidaner H, Fidaner C, Eser SY, Elbi H, Göker E. Psychometric properties of the WHOQOL-100 and WHOQOL-BREF. *Psikiyatr Psikol Psikofarmakol (3P) Derg*. 1999;7(Suppl 2):23-40.
- Zimet GD, Dahlem NW, Zimet SG, Farley GK. The multidimensional scale of perceived social support. *J Person Assessment*. 1988;52(1):30-41. doi:10.1207/s15327752jpa5201_2
- Eker D, Arkar H. Perceived social support: psychometric properties of the MSPSS in normal and pathological groups in a developing country. *Soc Psychiatr Psychiatr Epidemiol*. 1995;30:121-126. doi:10.1007/BF00802040
- Ritsher JB, Otilingam PG, Grajales M. Internalized stigma of mental illness: psychometric properties of a new measure. *Psychiatr Res*. 2003;121(1):31-49. doi:10.1016/j.psychres.2003.08.008
- Ersoy MA, Varan A. Reliability and validity of the Turkish version of the internalized stigma of Mental Illness Scale. *Turk J Psychiatr*. 2007;18(2):163-171.
- Şar V, Öztürk E, İkikardeş E. Validity and reliability of the Turkish version of childhood trauma questionnaire. *Turk Clin J Med Sci*. 2012;32(4):1054-1063. doi:10.5336/medsci.2011-26947
- Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. *Schizophr Bull*. 1987;13(2):261-276. doi:10.1093/schbul/13.2.261
- Kostakoğlu A, Batur S, Tiryaki A, Göğüş A. Reliability and validity of the Turkish version of the Positive and Negative Syndrome Scale. *Turk J Psychol*. 1999;14(44):23-34.
- Galuppi A, Turola MC, Nanni MG, Mazzoni P, Grassi L. Schizophrenia and quality of life: how important are symptoms and functioning? *Int J Mental Health Syst*. 2010;4:1-8. doi:10.1186/1752-4458-4-31
- İzci F, Fidan Y, Çallı S, Aslan M. Stigma, perception of social support, and functionality levels of schizophrenia patients with and without legal guardianship. *Psychiatr Clin Psychopharmacol*. 2024;34(1):74. doi:10.5152/pcp.2024.22602
- Fan C-H, Hsu S-C, Hsiao F-H, et al. The association of social support and symptomatic remission among community-dwelling schizophrenia patients: a cross-sectional study. *Int J Environment Res Public Health*. 2021;18(8):3977. doi:10.3390/ijerph18083977
- Munikanan T, Midin M, Daud TIM, et al. Association of social support and quality of life among people with schizophrenia receiving community psychiatric service: a cross-sectional study. *Compreh Psychiatr*. 2017;75:94-102. doi:10.1016/j.comppsy.2017.02.009
- Chan KKS, Fung WTW, Leung DCK, Tsui JKC. The impact of perceived and internalised stigma on clinical and functional recovery among people with mental illness. *Healt Social Care Commun*. 2022;30(6):e6102-e6111. doi:10.1111/hsc.14047
- Yanos PT, Lysaker PH, Silverstein SM, et al. A randomized-controlled trial of treatment for self-stigma among persons diagnosed with schizophrenia-spectrum disorders. *Soc Psychiatr Psychiatr Epidemiol*. 2019;54:1363-1378. doi:10.1007/s00127-019-01702-0
- Sarraf L, Lepage M, Sauve G. The clinical and psychosocial correlates of self-stigma among people with schizophrenia spectrum disorders across cultures: a systematic review and meta-analysis. *Schizophr Res*. 2022;248:64-78. doi:10.1016/j.schres.2022.08.001

39. Howell MK, Marggraf M, Taylor ML, et al. Social functioning mediates the relationship between psychiatric symptoms and recovery among veteran and community service users with serious mental illness. *Psychiatr Rehabil J*. 2023;46(2):156. doi:10.1037/prj0000559
40. Starzer M, Hansen HG, Hjorthøj C, Albert N, Glenthøj LB, Nordentoft M. Long-term quality of life and social disconnection 20 years after a first episode psychosis, results from the 20-year follow-up of the OPUS trial. *Schizophr Res*. 2024;274:33-45. doi:10.1016/j.schres.2024.08.021
41. Ercan Doğu S, Örsel S. The relationship between psychopathology, occupational balance, and quality of life among people with schizophrenia. *Austral Occupat Ther J*. 2023;70(3):314-326. doi:10.1111/1440-1630.12855
42. Grover S, Sahoo S, Chakrabarti S, Avasthi A. Association of internalized stigma and insight in patients with schizophrenia. *Int J Culture Mental Health*. 2018;11(3):338-350. doi:10.1080/17542863.2017.1381750
43. Tworek G, Thompson NR, Kane A, Sullivan AB. The impact of stigma on perceived quality of life and experience of anxiety and depression in individuals diagnosed with MS. *Multiple Sclerosis Relat Disord*. 2023;72:104591. doi:10.1016/j.msard.2023.104591
44. Vass V, Morrison AP, Law H, et al. How stigma impacts on people with psychosis: the mediating effect of self-esteem and hopelessness on subjective recovery and psychotic experiences. *Psychiatr Res*. 2015;230(2):487-495. doi:10.1016/j.psychres.2015.09.042
45. Kilian S, Asmal L, Phahladira L, et al. The association between childhood trauma and treatment outcomes in schizophrenia spectrum disorders. *Psychiatr Res*. 2020;289:113004. doi:10.1016/j.psychres.2020.113004
46. Inyang B, Gondal FJ, Abah GA, et al. The role of childhood trauma in psychosis and schizophrenia: a systematic review. *Cureus*. 2022;14(1):e21466. doi:10.7759/cureus.21466
47. Alkema A, Marchi M, Van Der Zaag JA, et al. Childhood abuse v. neglect and risk for major psychiatric disorders. *Psychol Med*. 2024;54(8):1598-1609. doi:10.1017/S0033291723003471
48. Teicher MH, Samson JA, Anderson CM, Ohashi K. The effects of childhood maltreatment on brain structure, function and connectivity. *Nat Rev Neurosci*. 2016;17(10):652-666. doi:10.1038/nrn.2016.111
49. Burger SR, van der Linden T, Hardy A, et al. Trauma-focused therapies for post-traumatic stress in psychosis: study protocol for the RE-PROCESS randomized controlled trial. *Trials*. 2022;23(1):851. doi:10.1186/s13063-022-06808-6

Global research trends and collaborative networks in renal colic: a bibliometric analysis of the 100 most cited articles

Çağrı Öktem¹, Veysi Siber²

¹Department of Urology, Kulu State Hospital, Konya, Türkiye

²Department of Emergency Medicine, Ankara Etlik City Hospital, Ankara, Türkiye

Cite this article as: Öktem Ç, Siber V. Global research trends and collaborative networks in renal colic: a bibliometric analysis of the 100 most cited articles. *J Health Sci Med.* 2025;8(6):975-980.

Received: 22.08.2025

Accepted: 13.09.2025

Published: 25.10.2025

ABSTRACT

Aims: Renal colic is a prevalent urological emergency, yet the global research landscape on this condition has not been systematically evaluated. Bibliometric analysis provides insights into scientific productivity, citation impact, collaboration patterns, and emerging research trends. The aim of this study was to evaluate scientific productivity, citation impact, and collaboration networks in the field of renal colic by analyzing the 100 most-cited original research articles.

Methods: A bibliometric study was conducted using the Web of Science Core Collection database. Following inclusion and exclusion criteria, the 100 most-cited original research articles on renal colic were identified. Data cleaning was performed with BibExcel, and co-authorship, co-occurrence, and bibliographic coupling analyses were conducted using VOSviewer (version 1.6.20).

Results: A total of 3367 eligible publications were retrieved, from which the 100 most-cited original articles were analyzed. These articles were published in 33 journals, authored by 508 researchers from 26 countries, and accumulated significant citation impact. The United States emerged as the most productive and influential country (h-index 46), with Yale University leading at the institutional level. The most impactful journals included American Journal of Roentgenology, Journal of Urology, and Radiology. Prominent authors were Smith RC, Verga M, and Rosenfield AT, all of whom demonstrated high total link strength. Keyword analysis revealed that “renal colic,” “computed tomography,” “radiation dose,” and “calculi” were the most frequent and central research themes, while bibliographic coupling highlighted clusters emphasizing diagnostic imaging and radiation safety.

Conclusion: This bibliometric analysis identified the United States as the most productive country, Yale University as the leading institution, and American Journal of Roentgenology, Journal of Urology, and Radiology as the most influential journals. Diagnostic imaging and radiation safety emerged as central research themes. Strengthening international collaboration and generating data from low- and middle-income countries are critical to addressing current gaps.

Keywords: Renal colic, bibliometric, citation, research, tomography

INTRODUCTION

Renal colic is one of the most common urological emergencies, characterized by severe flank pain resulting from acute obstruction caused by ureteral calculi.¹ Its clinical spectrum ranges from mild discomfort to excruciating pain, and emergency department visits due to acute colic are frequent. The high rate of recurrent presentations, particularly among young and middle-aged adults, coupled with the need for advanced imaging modalities such as computed tomography (CT) or ultrasound in diagnostic and therapeutic processes, renders renal colic not only a significant economic burden on healthcare systems but also a clinically critical condition for physicians.²⁻⁴

In recent years, significant changes have occurred in the diagnostic approach within urology and emergency medicine practice. Advances in the diagnostic sensitivity of CT the widespread adoption of low-dose protocols, ongoing debates

regarding the reduction of radiation exposure, and the evaluation of conservative versus interventional treatment options based on stone size have emerged as prominent areas of research in the literature.⁵ Moreover, epidemiological data, variations in prevalence rates across countries, and environmental risk factors have further contributed to the growing scientific interest in renal colic.⁶

Bibliometric analyses provide a systematic means of evaluating scientific output, citation impact, and collaboration networks within a specific field. Such analyses not only identify the most highly cited studies but also highlight research trends, levels of collaboration across countries and institutions, as well as leading journals and authors, thereby guiding future investigations.⁷ In particular, determining which centers produce high-level evidence that influences clinical decision-making, identifying the methodologies that distinguish these

Corresponding Author: Çağrı Öktem, cagrioktemgu@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

studies, and clarifying the topics they emphasize play a critical role in revealing knowledge gaps within the literature.⁸

The present study aims to evaluate scientific productivity, citation impact, and collaboration networks in the field of renal colic through a bibliometric analysis of the literature indexed in the Web of Science Core Collection. To this end, the 100 most-cited original research articles were analyzed, providing a comprehensive assessment in terms of contributing countries, institutions, authors, journals, and keyword trends. This study seeks to offer an integrative perspective on the historical evolution, current landscape, and future research directions of renal colic investigations. By systematically mapping the most-cited literature, this study contributes to the scientific field by identifying leading countries, institutions, and research themes, while also clarifying existing knowledge gaps and guiding future investigations in renal colic.

METHODS

Ethics

As this study is a bibliometric analysis based on data extracted from the Web of Science Core Collection, it did not involve human participants, patient data, or animal experiments; therefore, ethics committee approval was not required. All procedures were carried out in accordance with the ethical rules and the principles.

Search Strategy and Data Collection

The Web of Science (WoS) platform (<https://www.webofscience.com/>) is a subscription-based research system that provides extensive citation data across multidisciplinary academic fields. For the present study, bibliometric analysis was conducted using the WoS Core Collection database. The bibliometric search covered publications from 1980 through 2024. All data retrieval procedures were performed on April 23, 2025, ensuring that the analysis included the complete set of records available for the year 2024. During the data collection process, the advanced search mode of WoS was utilized, applying the “topic” (TS) field with the following keyword combination: TS=(“renal colic” OR “ureter* colic” OR “acute renal colic” OR “acute ureter* colic” OR “flank pain” OR “renal pain”).

Inclusion Criteria:

- Articles published in journals indexed within the Science Citation Index Expanded (SCIE) of WoS
- Studies published exclusively in English
- Original research articles (including both open-access and subscription-based publications)
- Publications retrieved using the predefined keyword search strategy

Exclusion Criteria:

- Publication types other than original research articles (e.g., case reports, case series, reviews, systematic reviews, meta-analyses, etc.)

Based on the predefined inclusion and exclusion criteria, the 100 most-cited original research articles were retrieved

from the WoS Core Collection database. The records were downloaded in text file format, and visualizations, mapping, and bibliometric analyses were subsequently performed using various data analysis tools.

Data Cleaning

A comprehensive data-cleaning process was performed prior to analysis. In the initial stage, spelling inconsistencies in author names, country names, and journal titles extracted from the WoS database were manually corrected. Subsequently, BibExcel was used to merge different variants of the same author or country, thereby eliminating duplicate records. Finally, a thesaurus file was applied in VOSviewer analyses to ensure consistency, consolidating all variations of author names into a single standardized format. These steps were undertaken to enhance both the reliability of the analyses and the accuracy of the visual outputs.

Statistical Analysis

Data organization and table construction were performed using Microsoft Excel (version 2016).⁹ For processing bibliometric data and identifying the most productive authors, countries, and journals, BibExcel software (version 2016-02-20) was utilized.¹⁰ For each category, the number of articles, total citation count, h-index, and citation sum within the h-core were calculated. Additionally, the average citation count was computed for the most productive journals. Although the initial plan was to list only the top 10 ranks, all authors or countries with equal h-index values were included in the tables (for example, 11 authors were evaluated due to identical h-index scores).

Visual maps were generated using VOSviewer (version 1.6.20).¹¹ The analyses employed co-authorship, co-occurrence, and bibliographic coupling methods. In all analyses, the full counting approach was adopted, while the association strength method was applied for calculating link strength. Parameter settings included random start=1000 and maximum iterations=1000, aiming to achieve optimal layout from different starting points.

In the co-authorship analysis, collaboration between authors and institutions was evaluated. The link length was inversely proportional to the strength of the relationship, while line thickness reflected the intensity of collaboration. In the author co-authorship analysis, each node represented an author, and the connecting lines indicated joint publications. The analysis was conducted with a minimum document threshold of 2 and a minimum citation threshold of 0; disconnected nodes were also included in the visualization. Node sizes were proportional to the total link strength (TLS). In the institutional co-authorship analysis, each node represented an institution, and the connecting lines reflected collaborative works. For this analysis, the minimum document threshold was set to 1 and the minimum citation threshold to 0; the largest set of connected items was extracted and visualized.

In the keyword co-occurrence analysis, the frequency of simultaneous appearance of author-reported keywords in the articles was examined. In the overlay visualization, colors represented temporal trends in keyword usage, with newer

terms displayed in yellow and older terms in purple. The network visualization demonstrated the associations between the most frequently used keywords and related terms. The inclusion threshold for the analysis was set at a minimum occurrence of one.

The bibliographic coupling analysis was conducted to identify structural linkages within the literature and to cluster publications with similar citation patterns. A minimum citation threshold of zero was applied, and citation counts were used to weight the strength of the connections. In the visualization, node sizes represented the number of citations, while colors indicated clusters of publications sharing similar citation profiles.

RESULTS

Applying the predefined search strategy for renal colic in the Web of Science database yielded a total of 6,731 publications. After applying the inclusion and exclusion criteria, 3,367 publications remained. From these, the 100 most-cited original research articles were identified. A total of 508 authors from 26 countries contributed to these articles, which were published across 33 journals (Figure 1).

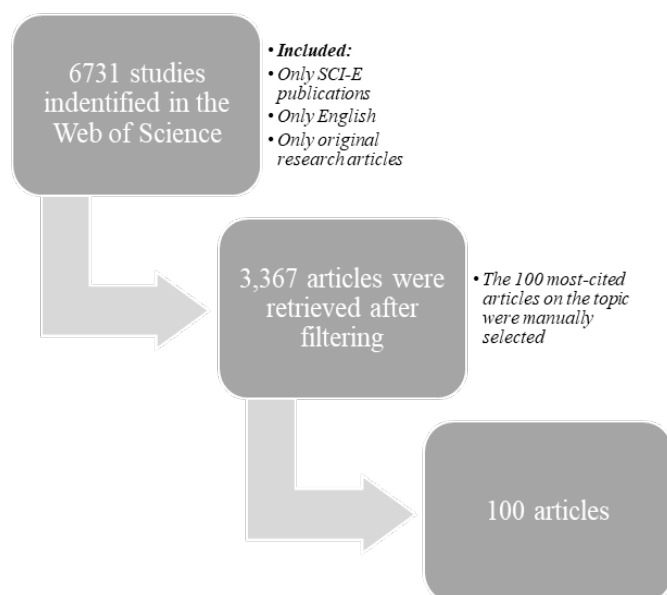


Figure 1. Flow chart of the inclusion process

Table 1. The most productive and effective countries in the 100 most cited articles on renal colic

No	Countries	All articles	Citation sum within h-core	h-index
1	USA	46	3993	46
2	Canada	5	646	5
3	Turkiye	5	455	5
4	Belgium	4	421	4
5	UK	4	468	4
6	Finland	3	172	3
7	Germany	3	793	3
8	Switzerland	3	425	3
9	Italy	2	146	2
10	Peoples R China	2	139	2
11	Saudi Arabia	2	125	2

The countries with the highest h-index (h-i) were the USA (46 h-i), Canada (5 h-i), Turkiye (5 h-i), Belgium (4 h-i), and the UK (4 h-i). In terms of citation sum within the h-core (CSh), the leading countries were the USA (46 articles, 3,993 CSh), Germany (3 articles, 793 CSh), and Canada (5 articles, 646 CSh) (Table 1).

When considering both citation impact and productivity, the journals with the highest h-index values were American Journal of Roentgenology (15 h-i, 2,069 CSh), Journal of Urology (13 h-i, 1,413 CSh), and Radiology (12 h-i, 1,629 CSh). The journals with the highest average citation counts (AvgCit) were New England Journal of Medicine (1 h-i, 271.00 AvgCit), European Urology (6 h-i, 153.00 AvgCit), and The Lancet (3 h-i, 144.67 AvgCit) (Table 2).

The authors with the highest h-index and corresponding citation sum within the h-core (CSh) were Smith RC (7 h-i, 1,682 CSh), Verga M (6 h-i, 1,606 CSh), and Rosenfield AT (6 h-i, 1,506 CSh) (Table 3).

Figure 2 presents the network map illustrating international collaboration among countries. Among the 508 authors, a co-authorship analysis was conducted for those with at least two publications and a minimum of zero citations, yielding a total of 33 authors. Total link strength (TLS) was selected as the weighting unit. While some nodes were unconnected,

Table 2. The most productive and effective journals in the 100 most cited articles on renal colic

No	Journal names	All articles	Citation sum within h-core	h-index	Average citation
1	American Journal of Roentgenology	15	2069	15	137,93
2	Journal of Urology	13	1413	13	108,69
3	Radiology	12	1629	12	135,75
4	Annals of Emergency Medicine	7	497	7	71,00
5	European Urology	6	918	6	153,00
6	Urology	4	414	4	103,50
7	Academic Emergency Medicine	3	186	3	62,00
8	Lancet	3	434	3	144,67
9	Clinical Radiology	3	189	3	63,00
10	European Radiology	3	408	3	136,00
11	BJU International	3	144	3	48,00

Table 3. The most productive and effective authors in the 100 most cited articles on renal colic

No	Authors	All articles	Citation sum within h-core	h-index
1	Smith, RC	7	1682	7
2	Verga, M	6	1606	6
3	Rosenfield, AT	6	1506	6
4	Dalrymple, NC	4	685	4
5	Moore, CL	4	257	4
6	Chen MYM	3	229	3
7	Zagoria, RJ	3	229	3
8	Daniels, B	3	162	3
9	Luty, Seth	3	162	3
10	Molinaro, AM	3	162	3
11	Singh, Dinesh	3	162	3

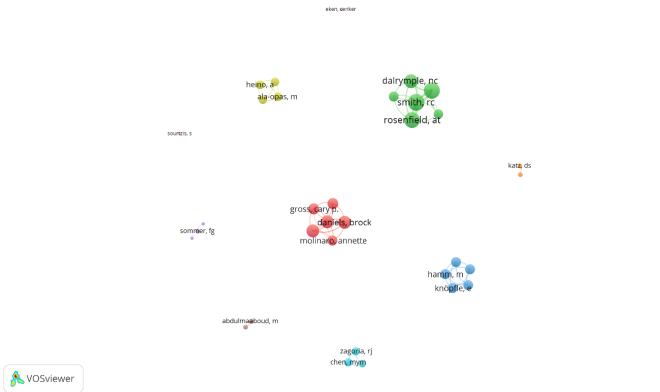


Figure 2. Network visualization map of co-authorships among authors in the 100 most cited publications in the field of renal colic

the largest connected component consisted of six items; nevertheless, the visualization included all eligible nodes to reflect the overall structure of the network. The analysis identified 10 clusters, 52 links, and a cumulative TLS of 119. Within this co-authorship network, the authors with the highest total link strength, as well as the greatest number of documents and citations, were Smith RC (20 TLS, 7 articles, 1,682 citations), Verga M (19 TLS, 6 articles, 1,606 citations), Rosenfield AT (18 TLS, 6 articles, 1,506 citations), and Dalrymple NC (14 TLS, 4 articles, 685 citations).

Figure 3 presents the network map of institutional collaboration. A co-authorship analysis was conducted among 146 institutions that had contributed at least one publication and received a minimum of zero citations, with documents selected as the weighting unit. While some nodes were unconnected, the largest connected component consisted of 25 institutions, and the analysis was therefore performed within this cluster. Overall, the analysis identified four clusters, 121 links, and a total link strength (TLS) of 123. When assessed in terms of documents, citations, and TLS, Yale University emerged as the most influential institution. The top five institutions by publication output were Yale University (10 articles, 1,031 citations, 24 TLS), Duke University (5

articles, 481 citations, 4 TLS), Wake Forest University (4 articles, 369 citations, 14 TLS), Akdeniz University (3 articles, 338 citations, 2 TLS), and the University of California, San Francisco (3 articles, 162 citations, 4 TLS).

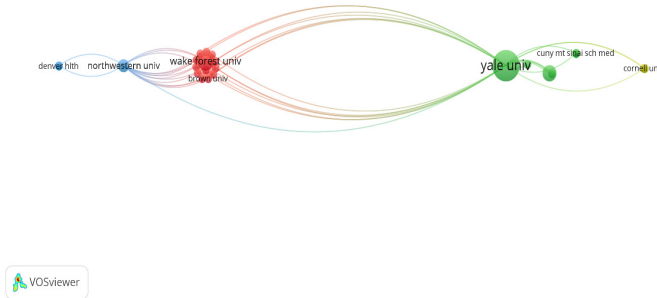


Figure 3. Network visualization map of co-authorships among organizations in the 100 most cited publications in the field of renal colic

Figure 4 illustrates the network of connections among author keywords over time. In the overlay visualization, nodes shaded toward yellow indicate more recent publications, while those shaded toward purple represent earlier studies. The minimum occurrence threshold for inclusion was set at one document. A total of 170 keywords were identified, of which 154 exhibited interconnections and were included in the co-occurrence overlay visualization. Node weighting was based on occurrences, and the score value applied in the overlay map was the average publication year (APY), enabling visualization of the temporal trends in keyword usage. The analysis revealed 154 connected nodes, 515 links, and a total link strength (TLS) of 560. Overall, strong keyword connections were observed between 1998 and 2007. The most frequently co-occurring keywords were renal colic (13 occurrences, 30 TLS, 44 links, APY 2004.46), radiation dose (7 occurrences, 30 TLS, 20 links, APY 2003.29), calculi (7 occurrences, 27 TLS, 21 links, APY 2001.29), CT (7 occurrences, 26 TLS, 18 links, APY 2005.00), ureter (6 occurrences, 29 TLS, 20 links, APY 2001.00), urography (6 occurrences, 29 TLS, 24 links, APY 1998.83), and computed tomography (6 occurrences, 28 TLS, 24 links, APY 2006.67).

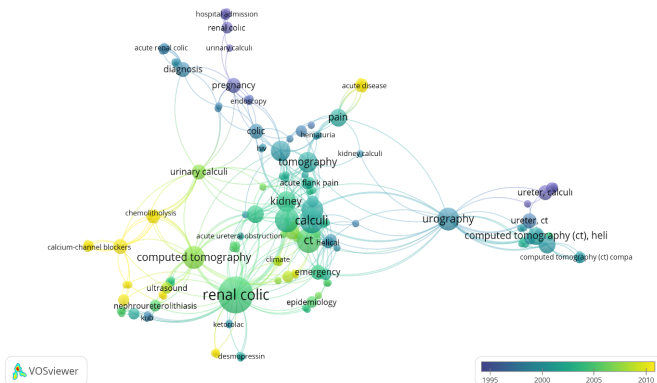


Figure 4. The author keywords identified in the 100 most cited articles on renal colic were subjected to co-occurrence analysis, and their relationships were mapped using overlay visualization

Figure 5 presents the bibliographic coupling analysis conducted to identify relationships among documents with similar citation patterns. Among the top 100 most-cited articles on renal colic, 98 demonstrated interconnections and were visualized in the network map. The analysis revealed six clusters, comprising 1,523 links and a total link strength (TLS) of 4,643. The documents with the highest citation counts were Tuerk (2016) (612 citations, 39 TLS), Smith (1995) (507 citations, 16 TLS), Smith (1996b) (414 citations, 64 TLS), Teichman (2004) (271 citations, 67 TLS), and Poletti (2007) (264 citations, 165 TLS). In contrast, the documents with the highest overall link strength, reflecting the strongest structural connections within the network, were Coursey (2012) (140 citations, 321 TLS), Memarsadeghi (2005) (61 citations, 300 TLS), Catalano (2002) (97 citations, 255 TLS), Pfister (2003) (91 citations, 255 TLS), and Ege (2003) (46 citations, 250 TLS).

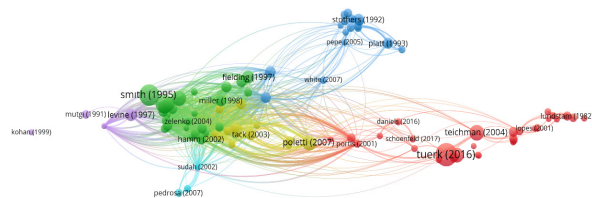


Figure 5. A bibliographic coupling analysis was conducted on the 100 most cited documents in the field of renal colic

DISCUSSION

This study provides a comprehensive bibliometric evaluation of the top 100 most-cited original research articles in the field of renal colic, focusing on scientific productivity, citation impact, collaboration networks, and research trends. The findings highlight both the historical evolution and the current dynamics of research in this domain, while also identifying existing knowledge gaps and potential directions for future investigations.

According to the results, the United States emerged as by far the most productive country with the highest citation impact. This finding is consistent with trends previously reported in bibliometric studies on urolithiasis and urinary stones.^{12,13} At the institutional level, Yale University ranked first in terms of publication count, citation impact, and total link strength, underscoring the pivotal role of U.S.-based academic centers in shaping the research agenda in this field.

Among the journals with the highest h-index, the American Journal of Roentgenology, Journal of Urology, and Radiology were the most prominent, while the New England Journal of Medicine, European Urology, and The Lancet demonstrated the highest average citation counts. These findings indicate that research on renal colic is not confined to urology and radiology journals but also exerts substantial influence in multidisciplinary outlets. Notably, studies focusing on the advancement of CT utilization and the development of low-dose protocols have gained particular prominence within the radiology and emergency medicine literature.¹⁴

Among the most productive and influential authors were Smith RC, Verga M, and Rosenfield AT. The fact that these authors also demonstrated high total link strength underscores the role of collaboration networks in enhancing bibliometric impact. Guo et al.,¹⁵ in their bibliometric analysis of ureteroscopy, similarly reported that core authors exhibited strong link strength, thereby increasing the visibility of collaborative networks. Likewise, Fagan et al.¹⁶ analyzed co-authorship data through social network analysis and noted the formation of strong collaborations within specific centers, while many authors and institutions remained isolated. These findings are consistent with the results of our study, highlighting that scientific output in the field tends to concentrate within certain hubs, whereas international collaborations remain relatively limited.

The keyword analysis highlighted “renal colic,” “computed tomography,” “radiation dose,” and “calculi” as dominant terms, underscoring the central role of diagnostic imaging and radiation safety in this field. Recent studies have also pointed toward emerging trends such as ultra-low-dose CT protocols and the increasing application of point-of-care ultrasound, which not only improve diagnostic accuracy but also address health economics and patient safety. These trends suggest a shift in renal colic research toward optimizing diagnostic strategies, balancing accuracy with safety, and exploring cost-effective approaches suitable for diverse healthcare systems.^{17,18}

The findings of this study highlight several critical gaps in the existing literature. First, the majority of highly cited studies are concentrated in North America and Western Europe, with a markedly limited contribution from low- and middle-income countries. Brück emphasizes the “Anglocentric” nature of medical publishing and the underrepresentation of developing countries in the global research landscape.¹⁹ This imbalance reveals a paucity of epidemiological data from regions where environmental risk factors such as climate, dietary habits, and fluid intake differ substantially yet remain underexplored.²⁰ Furthermore, both bibliometric and epidemiological analyses confirm that this disparity contributes to a global imbalance in knowledge dissemination.²¹ Second, while the most frequently cited studies predominantly focus on diagnostic imaging modalities, there is a relative paucity of research addressing treatment algorithms, pharmacologic strategies, and patient-centered outcomes. This observation underscores a significant gap in the literature concerning conservative management, interventional decision-making based on stone size, and long-term follow-up data.

Limitations

Among the strengths of the present study are the exclusive inclusion of original research articles published in journals indexed in the SCIE, which ensured a high level of scientific rigor. Furthermore, the application of multiple bibliometric approaches—including co-authorship, co-occurrence, and bibliographic coupling analyses—enhanced the robustness and reliability of the findings. However, this study has certain limitations. First, the analysis was restricted to the Web of Science Core Collection database. While WoS is widely regarded as the standard source for bibliometric research due to its rigorous indexing criteria and citation tracking system,

the inclusion of Scopus could potentially capture a broader range of journals and provide complementary insights. Second, only articles published in English were included, which may have led to the omission of regionally published studies in other languages. This language restriction could exclude valuable research findings and perspectives, particularly from non-English speaking regions, thereby limiting the global representativeness of the results.

CONCLUSION

This bibliometric analysis revealed that the United States was the most productive and influential country, with Yale University ranking first at the institutional level. The American Journal of Roentgenology, Journal of Urology, and Radiology were identified as the most impactful journals, and Smith RC, Verga M, and Rosenfield AT as the most influential authors. Keyword analysis underscored diagnostic imaging and radiation safety as central research themes, while bibliographic coupling highlighted clusters around CT utilization and radiation dose optimization. Moving forward, fostering international collaborations through multi-center trials that include institutions from both high- and low-to-middle income countries and investigating cost-effective diagnostic and treatment pathways tailored for resource-limited settings, alongside promoting treatment-focused clinical studies, will be essential to addressing knowledge gaps and improving clinical decision-making in renal colic.

ETHICAL DECLARATIONS

Ethics Committee Approval

As this study is a bibliometric analysis based on data extracted from the Web of Science Core Collection, it did not involve human participants, patient data, or animal experiments; therefore, ethics committee approval was not required.

Informed Consent

Since this research is a bibliometric study, it did not require informed consent.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Szymanski J, Chłosta M, Dudek P, et al. Prevalence, correlates, and treatment behaviors for urolithiasis and renal colic-like pain symptoms at the population level in Poland. *Sci Rep*. 2025;15(1):10827. doi:10.1038/s41598-025-95504-x
2. Kepka S, Zarca K, Ohana M, et al. A real-world cost-effectiveness study evaluating imaging strategies for the diagnostic workup of renal colic in the emergency department. *Medicina (Kaunas)*. 2023;59(3):475. doi:10.3390/medicina59030475
3. Wang RC, Fahimi J, Dillon D, et al. Effect of an ultrasound-first clinical decision tool in emergency department patients with suspected nephrolithiasis: a randomized trial. *Am J Emerg Med*. 2022;60:164-170. doi:10.1016/j.ajem.2022.08.015
4. Bartoletti R, Cai T, Mondaini N, et al. Epidemiology and risk factors in urolithiasis. *Urol Int*. 2007;79(Suppl 1):3-7. doi:10.1159/000104434
5. Smith-Bindman R, Moghadassi M, Griffey RT, et al. Computed tomography radiation dose in patients with suspected urolithiasis. *JAMA Intern Med*. 2015;175(8):1413-1416. doi:10.1001/jamainternmed.2015.2697
6. GBD 2021 Urolithiasis Collaborators. The global, regional, and national burden of urolithiasis in 204 countries and territories, 2000-2021: a systematic analysis for the global burden of disease study 2021. *EClinicalMedicine*. 2024;78:102924. doi:10.1016/j.eclinm.2024.102924
7. Donthu N, Kumar S, Mukherjee D, Pandey N, Lim WM. How to conduct a bibliometric analysis: an overview and guidelines. *J Business Res*. 2021;133:285-296.
8. Martín-Martín A, Thelwall M, Orduna-Malea E, Delgado López-Cózar E. Google Scholar, Microsoft Academic, Scopus, Dimensions, Web of Science, and OpenCitations' COCI: a multidisciplinary comparison of coverage via citations. *Scientometrics*. 2021;126(1):871-906. doi:10.1007/s11192-020-03690-4
9. Microsoft Corporation. (2016). Microsoft Excel (Version 2016). 2016 [Available from: <https://office.microsoft.com/excel>]
10. Persson O, Danell R, Schneider JW. How to use Bibexcel for various types of bibliometric analysis. Celebrating scholarly communication studies: a Festschrift for Olle Persson at his 60th Birthday. 2009;5(2009):9-24.
11. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*. 2010;84(2):523-538. doi:10.1007/s11192-009-0146-3
12. Zhong Y, Zeng Q, Yi J, Long T, Peng J, Zhong L. Research trends and frontiers on risk factors of urinary stones: a bibliometric analysis from 2010 to 2023. *Ren Fail*. 2024;46(2):2440513. doi:10.1080/0886022X.2024.2440513
13. Abushamma F, Zyoud SH. Analyzing global research trends and focal points in the utilization of laser techniques for the treatment of urolithiasis from 1978 to 2022: visualization and bibliometric analysis. *Urolithiasis*. 2024;52(1):67. doi:10.1007/s00240-024-01568-1
14. Moore CL, Carpenter CR, Heilbrun ME, et al. Imaging in suspected renal colic: systematic review of the literature and multispecialty consensus. *Ann Emerg Med*. 2019;74(3):391-399. doi:10.1016/j.annemergmed.2019.04.021
15. Guo Z, Yang Y, Liu C. Knowledge-map analysis of ureteroscopy for urolithiasis. *J Multidiscip Healthc*. 2024;17:6001-6015. doi:10.2147/JMDH.S493581
16. Fagan J, Eddens KS, Dolly J, Vanderford NL, Weiss H, Levens JS. Assessing research collaboration through co-authorship network analysis. *J Res Adm*. 2018;49(1):76-99.
17. Zhang X, Zhang G, Xu L, et al. Application of deep learning reconstruction of ultra-low-dose abdominal CT in the diagnosis of renal calculi. *Insights Imaging*. 2022;13(1):163. doi:10.1186/s13244-022-01300-w
18. Wang S, Xia D, Zhang Z, et al. Mapping trends and hotspots regarding the use of ultrasound in emergency medicine: a bibliometric analysis of global research. *Front Public Health*. 2021;9:764642. doi:10.3389/fpubh.2021.764642
19. Brück O. A bibliometric analysis of geographic disparities in the authorship of leading medical journals. *Commun Med (Lond)*. 2023;3(1):178. doi:10.1038/s43856-023-00418-2
20. Dirie NI, Hassan J, Hussein AO, et al. Burden of urolithiasis: a systematic review of epidemiological and clinical trends in Somalia. *Afr J Urol*. 2024;30(1):64. doi:10.1186/s12301-024-00468-5
21. Borumandnia N, Fattahi P, Talebi A, et al. Longitudinal trend of urolithiasis incidence rates among world countries during past decades. *BMC Urol*. 2023;23(1):166. doi:10.1186/s12894-023-01336-0

Aberrant left hepatic artery arising from the left gastric artery: perioperative effects of a risky vascular anomaly in gastrectomy

✉ Hüseyin Kılavuz¹, ✉ Yiğit Düzköylü², ✉ Hüsnü Ozan Şevik¹, ✉ Oğuzhan Tekin¹,
✉ Aytül Hande Yardımcı³, ✉ Sercan Yüksel¹

¹Department of General Surgery, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye

²Department of Gastroenterological Surgery, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye

³Department of Clinical Radiology, Başakşehir Çam and Sakura City Hospital, İstanbul, Türkiye

Cite this article as: Kılavuz H, Düzköylü Y, Şevik HO, Tekin O, Yardımcı AH, Yüksel S. Aberrant left hepatic artery arising from the left gastric artery: perioperative effects of a risky vascular anomaly in gastrectomy. *J Health Sci Med.* 2025;8(6):981-987.

Received: 23.08.2025

Accepted: 14.09.2025

Published: 25.10.2025

ABSTRACT

Aims: The aim of this study was to analyze the results of the choice of cutting and preservation in different surgical methods in the presence of aberrant left hepatic artery (ALHA) during gastric cancer surgery.

Methods: This study was conducted as a single-center, retrospective study between 01/01/2022 and 31/08/2023. Patients who underwent open, laparoscopic and robotic surgery due to gastric cancer were included. The presence of ALHA in preoperative computed tomography scans was screened using the application focusing on vascular anatomy in the radiology automation system. The results of patients with and without ALHA were analyzed.

Results: No statistically significant difference was observed between the groups in terms of duration of surgery, length of hospital stay, history of neoadjuvant therapy, total number of lymph nodes removed, and number of metastatic lymph nodes.

Conclusion: We recommend close postoperative monitoring of liver enzymes and collaboration with an experienced radiologist in patients with suspected ALHA.

Keywords: Aberrant left hepatic artery, gastric cancer, oncological surgery, robotic surgery

INTRODUCTION

Gastric cancer is the fifth most common malignancy and the third leading cause of cancer-related deaths worldwide.^{1,2} Notwithstanding the recent advances in multidisciplinary approaches, including adjuvant treatment models, immunotherapy and targeted therapies, surgery with negative margins and adequate lymphadenectomy remains the gold standard curative treatment.^{3,4}

It is widely acknowledged that ligating the main vessels at their origin and meticulously dissecting the surrounding tissue, ensuring complete removal of the gastrohepatic ligament, is paramount for achieving oncological and adequate resection.³ However, vascular variations in the left gastric and hepatic arteries are frequently encountered, impeding procedural efficacy and giving rise to contentious debates. The prevalence of vascular variations in hepatic artery anatomy has been documented to range from 13%-48% in the general population and it has been further reported that 10% to 25% of these variations are classified as aberrant left hepatic artery (ALHA), which originates from the left gastric artery (LGA).^{5,6} ALHA can be classified into two main subtypes: accessory LHA (acLHA)

and replaced LHA (RLHA), with a ratio of 45% and 55%, respectively.⁷ ALHA supplies the left hepatic lobe in part or in full, and is usually explored coincidentally during surgery.^{8,9} Although open surgery may ensure the identification of ALHA by palpation of the lesser omentum, this is challenging in laparoscopic surgery or obese patients.⁹ Curative gastric cancer surgery with extended lymphadenectomy requires the ligation of LGA at its root for station 7th lymph node and all the tissues forming the lesser omentum, causing the division of ALHA if it is present.^{3,10} Despite the existence of reports recommending the preservation of ALHA to prevent fatal liver dysfunction and ischemia, there remains considerable controversy among surgeons regarding the decision to ligate for oncological reasons.^{4,11}

Shinohara et al.¹⁰ found that there were no differences in terms of adequacy of lymph node dissection and oncological results in patients whether ALHA was preserved or not. However, Oki et al.¹² suggested that ALHA should be preserved when encountered in laparoscopic gastrectomy. In 2022, Tao et al.¹³ reported in their systematic review that ligation of ALHA is

Corresponding Author: Hüseyin Kılavuz, drhuseyinkilavuz@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

not always without risk and that measurement of the diameter of the artery may be a useful guide to determining the presence of chronic liver disease.

To date, a consensus on the safety and/or necessity of routine ligation of ALHA during gastric cancer surgery has not been reached. The objective of this study was to ascertain the variations of celiac artery branches and the frequency of ALHA presence by means of a retrospective re-evaluation of the preoperative radiological examinations of patients who underwent gastric cancer surgery. The secondary objective was to evaluate the effects of ALHA-oriented approaches and surgical procedures on liver function tests, perioperative results and oncological surgical principles.

METHODS

Ethics approval and Study Design

Our study was approved by the Başakşehir Çam and Sakura City Hospital Clinical Researches Ethics Committee (Date: 22.11.2023, Decision No: 602). All procedures were performed in accordance with ethical regulations and the tenets of the Declaration of Helsinki, and informed consent was obtained from patients for procedures performed in the clinic and during their hospital stay. This single-center and retrospective study followed the “STrengthening the Reporting of OBservational studies in Epidemiology (STROBE)” guidelines.

Inclusion criteria: To identify patients who underwent surgery for gastric cancer between January 2022, and August 2023, the hospital automation system retrospectively scanned the surgical codes for “subtotal gastrectomy” and “total gastrectomy.” Patients who underwent preoperative contrast-enhanced triphasic abdominal computed tomography (CT) were included in the study.

Exclusion criteria: Patients who underwent emergency surgery due to bleeding, obstruction, or perforation, or who underwent palliative surgery without oncological principles, were excluded. Patients who received non-contrast radiological imaging or unavailable images were excluded.

After identifying the cases to be included in the study, the contents of the surgery performed, details with vascular anatomy if indicated, and information on whether or not ALHA was ligated in case of deviation were recorded from the operation notes (Figure 1). Another experienced surgeon and a radiologist not involved in these procedures were asked to assess the celiac vascular anatomy according to the preoperative CT in a blinded manner (without knowledge of the surgical procedures) and to record it according to the presence of ALHA. In case of disagreement regarding vascular anatomy, these cases were evaluated jointly by the researchers and a decision was made. The radiologist's assessment was made using software that allows digital highlighting of the vascular network in the radiology automation system (Figure 2). Computed tomography images were obtained using SCT-7800TC (Shimadzu Corporation, Kyoto, Japan) and analyzed with a three-dimensional reconstruction software integrated into the Picture Archiving and Communication System (PACS) in Digital Imaging and Communications in Medicine (DICOM) format Synapse 3D (Fujifilm Corporation, Tokyo,

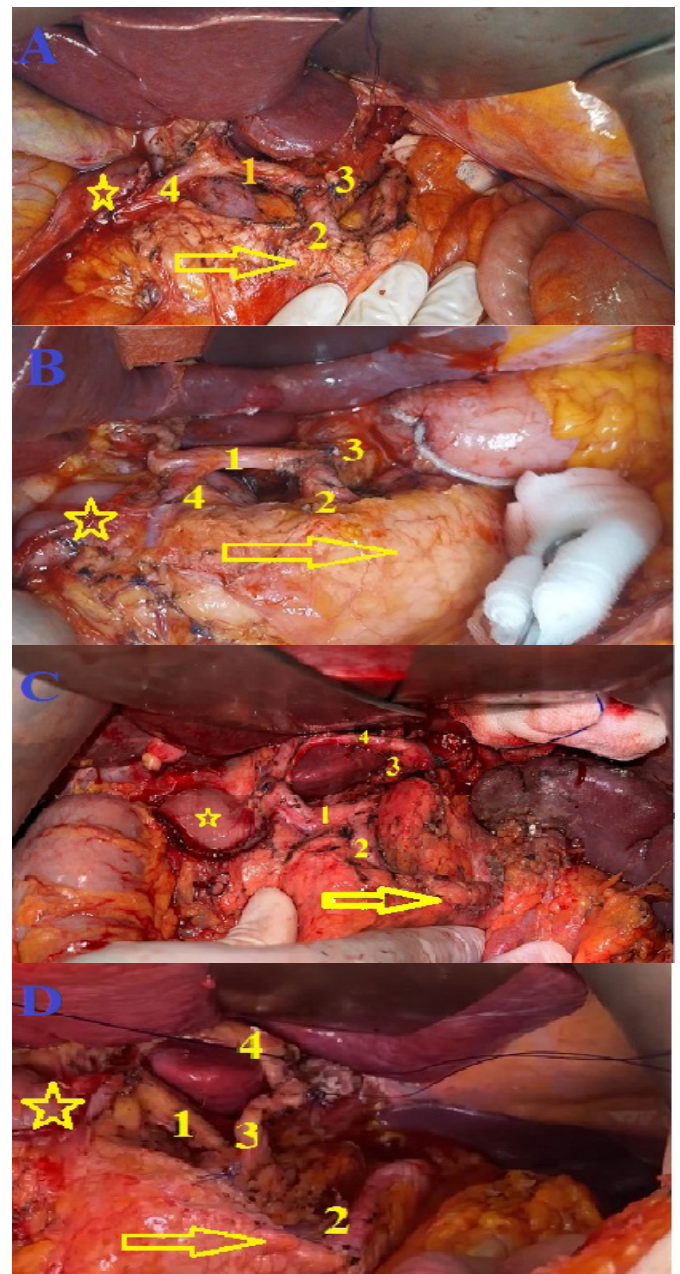


Figure 1. Examples of celiac arterial anatomy in our open surgeries. A) Total gastrectomy normal anatomy, B) Subtotal gastrectomy normal anatomy, C and D) examples of D2 lymph node dissection with preservation of the abbreviated left hepatic artery from different cases

Explanations of numbers and signs on the figure: 1) Common hepatic artery, 2) Splenic artery, 3) Left gastric artery, 4) Gastrooduodenal artery in A and B and 4) Aberrant left hepatic artery in C and D

Star: Duodenal stump and Arrow: Pancreas

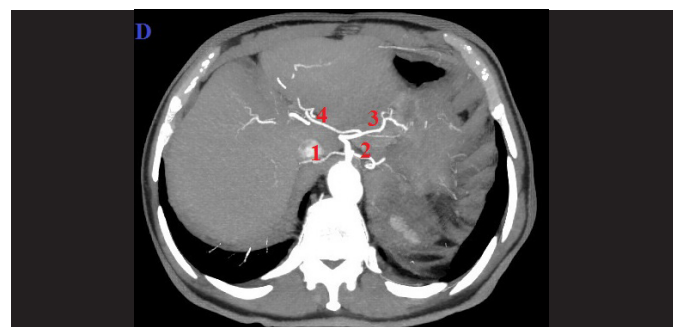


Figure 2. Demonstration of celiac vascular branches with radiology automation of the computerized tomography of the case in Figure 1D

Explanations of numbers on the figure: 1) Right hepatic artery, 2) Splenic artery, 3) Left gastric artery, 4) Aberrant left hepatic artery

Japan). Those without vascular variation were divided into two groups as group A and those with ALHA as group B. Group B was then recorded separately according to whether ALHA was cut or not.

Surgical Procedure

All patients are evaluated by a multidisciplinary gastrointestinal oncology surgery council. These councils include gastrointestinal surgeons, radiologists, medical oncologists, radiation oncologists, and pathologists. The council makes decisions regarding the suitability of surgical treatment, neoadjuvant therapy, or post-neoadjuvant surgery based on the clinical stage of gastric cancer. Patients who underwent surgery had a history of previous abdominal surgery, a high ASA (American Society of Anesthesiologists) score, and comorbidities that made laparoscopy/robotic surgery unsuitable for them underwent open surgery. A selection bias was applied to laparoscopic/robotic surgeries, favoring an experienced team. Younger patients, patients with a low body mass index, early-stage tumors without invasion of surrounding tissues, and surgical experience were the most important determining factors in this regard. In open surgery, a median incision was used, and all anastomoses were performed with staples. In laparoscopic procedures, anastomoses were performed intracorporally. The da Vinci Xi® Surgical System (Intuitive Surgical, Sunnyvale, CA, USA) was used as robotic technology.

Data Collection

The neoadjuvant treatment status of the cases, robotic/laparoscopic/open surgery status, total/subtotal gastrectomy status, operation time (minutes), hospital stay (days), total lymph node count and metastatic lymph node count in the pathology report were recorded. In laboratory tests, aspartate aminotransferase (AST, U/L), alanine aminotransferase (ALT, U/L) and total bilirubin (TB, U/L) levels were measured 48-72 hours before surgery and the results were recorded on the first, seventh and 30th postoperative days.

Statistical Analysis

SPSS 25.0 (SPSS Inc, Chicago, IL, USA) was used for statistical analysis of the data. While numerical variables in the data set were summarised by median (minimum-maximum), categorical variables were summarised by number (percentage). The Shapiro-Wilk test was used to test whether the data conformed to a normal distribution. In statistical analyses, the Mann-Whitney U test, the Friedman test and the Yates-corrected chi-square test were used, where appropriate,

given the distribution of the data. When statistically significant differences were found in the Friedman test, the Wilcoxon signed rank test with Bonferroni correction was used for pairwise comparisons. In statistical analyses, $p < 0.05$ was considered significant

RESULTS

A search of the database revealed that a total of 266 cases were operated on (open: 196, minimally invasive: 70) during the specified period. Of these, 209 cases met the inclusion criteria for the study. There were 173 patients in group A (no ALHA detected) and 36 patients in group B (ALHA detected). Of the total sample, 69 (33%) were female and 140 (67%) were male. The mean age of the sample was 63.7 ± 10.8 (26-86) years. The numbers of total/subtotal gastrectomy by robotic, laparoscopic and open methods were 43(20.5%)/14(6.7%), 14(6.7%)/7(3.3%) and 87(41.6%)/44(21%), respectively. According to celiac vascular anatomy, 173 (82.8%) cases with normal LGA were classified as group A and 36 (17.2%) cases with ALHA were classified as group B. Details of the groups are summarised in **Table 1**. Standard D2 lymph node dissection was performed in all cases in group A. D2 lymph node dissection was performed in all open cases in group B with preservation of ALHA (Figure 1), while lymphadenectomies were completed by ligation of the LGA root without preservation of ALHA in all robotic and laparoscopic cases. No statistically significant difference was observed in terms of operative time, days in hospital, history of neoadjuvant treatment, total number of lymph nodes removed and number of metastatic lymph nodes (**Table 2**).

Table 2. Comparison of surgical parameters of groups

	Group A	Group B	
Variables	Median (min-max)		P
Duration of surgery (minutes)	245(60-755)	242 (125-360)	0.792*
Duration of hospital stay (days)	7(3-41)	7(3-22)	0.673*
Pathology specimen total LN (n)	32(12-87)	32(17-75)	0.537*
Pathology specimen metastatic LN (n)	2(0-36)	3(0-40)	0.869*
Neoadjuvant treatment (yes)	77 (44.50)	21 (58.30)	0.184**
*: Mann-Whitney U test, **: Yatesin corrected Chi-square test, min-max: Minimum-maximum, LN: Lymph node			

In both groups, median values of laboratory parameters AST and ALT were elevated on the first postoperative day; however, these values decreased by the seventh postoperative

Table 1. Gastrectomy surgery methods and types

Group A (normal anatomy) n (%)												Group B (ALHA) n (%)			
173 (82.8)						36 (17.2)									
Total			Subtotal			Total			Subtotal						
118 (68.2)			55 (31.8)			26 (72.2)			10 (27.8)						
Robotic	Laparoscopic	Open	Robotic	Laparoscopic	Open	Robotic	Laparoscopic	Open	Robotic	Laparoscopic	Open				
35 (29.7)	11 (9.3)	72 (61)	14 (25.5)	7 (12.7)	34 (61.8)	8 (30.8)	3 (11.5)	15 (57.7)	0 (0)	0 (0)	10 (100)				
ALHA: Aberrant left hepatic artery															

ALHA: Aberrant left hepatic artery

day, approaching preoperative levels by the thirtieth day. Bilirubin levels remained unchanged throughout the postoperative period (Table 3). In both groups, a statistically significant difference was observed in the ALT values on the first, seventh and thirtieth postoperative days compared to the preoperative period ($p<0.05$). However, no significant difference was observed in the AST values of the groups on the thirtieth postoperative day (Table 4). In group B, no significant difference was observed between the postoperative AST and ALT averages in the subgroups with and without ALHA cut-off in terms of surgery duration, number of days of hospitalization, number of lymph nodes removed and number of metastatic lymph nodes (Table 5).

Table 3. Preoperative and postoperative liver function tests of the groups

	Group A	Group B
Variables	Median (min-max)	
ALT(U/L)		
Preoperative	14 (3-75)	12.5 (6-71)
Po1	48 (5-509)	50 (6-820)
Po7	26 (3-275)	21(6-195)
Po30	18 (3-141)	18 (3-374)
AST(U/L)		
Preoperative	20 (6-84)	18 (10-63)
Po1	50 (10-676)	55.5 (12-540)
Po7	26 (8-220)	25 (0.52-84)
Po30	22 (4-227)	20 (9-177)
Total bilirubin (U/L)		
Preoperative	0.33 (0.06-1.71)	0.31 (0.09-1.35)
Po1	0.51 (0.14-1.97)	0.445 (0.21-1.54)
Po7	0.44 (0.12-1.25)	0.41 (0.13-1.37)
Po30	0.38 (0.14-1.5)	0.41 (0.17-9.5)
min-max: Minimum-maximum, ALT: Alanine aminotransferase, AST: Aspartate aminotransferase Po: Postoperative day		

Table 4. Comparison of preoperative and postoperative liver function tests of the groups

Variable pairs		p-value for pairwise comparisons**	Overall p-value*
Group A	PreO ALT-po1ALT	<0.001	<0.001
	PreO ALT-po7ALT	<0.001	
	PreO ALT-po30ALT	0.002	
Group B	PreO ALT-po1ALT	<0.001	<0.001
	PreO ALT-po7ALT	<0.001	
	PreO ALT-po30ALT	0.018	
Group A	PreO AST-po1AST	<0.001	<0.001
	PreO AST-po7AST	<0.001	
	PreO AST-po30AST	0.106	
Group B	PreO AST-po1AST	<0.001	<0.001
	PreO AST-po7AST	0.407	
	PreO AST-po30AST	0.999	

ALT (U/L): Alanine aminotransferase, AST (U/L): Aspartate aminotransferase, PreO: Preoperative, Po: Postoperative day, *: Friedman test, **: Wilcoxon test with Bonferroni correction

Table 5. Analysis of group B variables

Group B variables	ALHA cuts (n=11)	ALHA protected (n=25)	p-value*
	Mean (±standard deviation)		
Preoperative ALT	11±6.06	13±13.23	0.918
Preoperative AST	22±14.19	17±9.15	0.126
po1 ALT	64±39.20	38±175.57	0.419
po1 AST	74±50.03	46±121.87	0.164
po7 ALT	17±7.63	25±39.00	0.068
po7 AST	19±7.52	26±20.40	0.074
po30 ALT	20±6.84	17±71.97	0.655
po30 AST	18±7.12	21±40.15	0.262
Duration of surgery (minutes)	240±64.15	245±60.38	0.439
Duration of hospital stay (days)	5±2.76	7±5.03	0.194
Pathology specimen total LN (n)	34±13.71	32±14.83	0.327
Pathology specimen metastatic LN (n)	2±2.75	4±11.76	0.058

ALHA: Aberrant left hepatic artery, ALT (U/L): Alanine aminotransferase, AST (U/L): Aspartate aminotransferase, LN: lymph node, Po: Postoperative day, *: Mann-Whitney U test

DISCUSSION

In our study, we retrospectively classified gastric cancer patients who underwent surgery into groups with and without ALHA to determine whether or not peroperative ligation of this artery is safe for liver function. Our results showed that 17.2% of the cases had an aberrant artery. Postoperative ALT levels were found to be elevated in both groups with and without ALHA, but no statistical difference was found at 30 days postoperatively. In the ALHA group, ligation or protection of the artery had no significant effect on postoperative AST-ALT levels, operation time, hospital stay, number of lymph nodes removed or number of metastatic lymph nodes.

Today, there is still no widely accepted consensus on ligating an ALHA during gastrectomy, especially if it is found incidentally. The gold standard curative treatment for gastric cancer is known to be total/subtotal gastrectomy with en-bloc lymphadenectomy, and regional lymph nodes circulating the root of the LGA should be dissected.³ For oncological safety, LGA should be ligated at its origin with en-bloc resection of lymph nodes three and six. However, an unrecognised or sacrificed left hepatic artery can lead to liver dysfunction and even death.

ALHA is the most common vascular variation encountered in gastric cancer surgery with rates ranging from 6.5% to 34%.¹⁴ Huang et al.¹¹ found the same anatomical variation with an incidence of 11.5% in their large retrospective analysis of 1173 patients. Essentially, ALHA can be divided into two main subtypes: accessory LHA (acLHA), which is an addition to the normal left hepatic artery, and replaced LHA (RLHA), which can be accepted as a replacement for a normal left hepatic artery.⁷ It is usually discovered incidentally during gastric cancer surgery due to the lack of a non-invasive and reliable preoperative imaging method.¹⁴ Liu et al.¹⁵ proposed preoperative three-dimensional CT angiography as a diagnostic imaging modality for the perigastric arterial system, which may not be cost-effective if routinely used in

the clinic. In our study, the celiac trunk and its branches were evaluated in detail by clarifying the contrast pattern of triphasic CT and arterial phase images thanks to the simulator on the digital system in radiology automation.

Although there are studies with contrary results, in general, the published literature usually recommends the preservation of ALHA to reduce postoperative complications. In their propensity score matching analysis, Lee et al.¹⁴ suggested that ALHA should be preserved, especially when it is thought to be replaced for postoperative liver function, and sacrificed when it is accessory for oncological safety. When the root of the LGA is ligated during lymphadenectomy for gastric cancer, the blood supply to the ALHA, if present, is also cut off. In this situation, fatal complications such as postoperative hepatic dysfunction, liver failure and left hepatic lobe necrosis due to ALHA ligation have been reported.¹⁶ Conversely, in their study of ligated and protected groups in both accessory and replacement ALHA patients, researchers found no differences in operative time, blood loss, overall complication rate, length of hospital stay and number of lymph nodes removed. It has been shown that if the ALHA diameter is greater than 1.5 mm, a transient increase in AST and ALT is observed on the 2nd postoperative day in ligated cases.⁸ Two weeks after surgery, all values recovered completely without the need for medication or intervention, similar to our results at 30 days.

The difference between laparoscopic and open surgery in relation with the presence of ALHA and liver function has been searched, focusing on the use of carbon dioxide to create pneumoperitoneum. Changes in liver function tests may be observed due to the intra-abdominal pressure resulting from pneumoperitoneum. Jeong et al.¹⁷ demonstrated this in their study by finding that AST and ALT levels were significantly higher on the first day after laparoscopic gastrectomy than after open gastrectomy. In addition to comparing open and laparoscopic surgery, the researchers also looked at whether the surgery was total or subtotal. In gastric cancer patients who underwent total or distal subtotal open gastrectomy, postoperative liver segments II and III were examined. The segments were found to have significant volume loss compared to the rest of the liver and the whole liver. This situation was associated with venous ischaemia due to the deterioration of the vessels contributing to the venous circulation of the gastric antrum.¹⁸ In our results, minimally invasive type or resection margins (total/subtotal) did not significantly affect on the results or found related with ALHA ligation.

In their multi-centre and comprehensive analysis, Ang et al. concluded that ALHA ligation is a safe procedure with no adverse outcomes, resulting only in a transient increase in post-operative AST-ALT levels in cases with more than 1.5 mm diameter of aberrant artery.⁸ As a result, the investigators suggested that ALHA be preserved in early gastric cancer and sacrificed in advanced cases. Pre-existing liver disease is also seen as a determining factor in the decision to withhold ALHA. Huang et al.¹¹ reported that it should not be ligated in this patient group to prevent liver dysfunction. Conversely, hepatic necrosis has not been reported even in patients with diagnosed liver disease after ALHA ligation, demonstrating that collateral arterial flow develops after hepatic artery ligation.^{10,13,14} Cirocchi et al.⁷ reported that collateralisation

between phrenic and intrahepatic arteries spares the left lobe of the liver from the first postoperative ten hours after ligation, explaining the transient increase in enzymes. Similarly, Kim et al.⁹ found no significant differences in operative time, blood loss, number of lymph nodes harvested or length of hospital stay after ALHA ligation. They also found no lethal complications such as necrosis or extreme increases in liver enzymes. The researchers also hypothesised that large ALHA result from a large LGA and that measuring the diameter of the LGA may be more accurate and easier than analysing ALHA preoperatively with contrast-enhanced CT. Lee et al.¹⁴ reported ALHA of 17% in their series and although the patients with postoperative elevated AST-ALT levels showed longer hospital stay and higher incidence of postoperative complications grade $\geq$ IIIa, they suggested sacrificing ALHA as a safe option. Interestingly, they also reported that only 20.8% of patients in the ligation group had a documented diagnosis of ALHA. Sano et al.¹⁶ found that ALHA ligation was a significant independent risk factor for liver enzyme elevation and suggested that ALHA should be preserved in the setting of oncologically standard dissection of the surrounding lymph nodes. In one of the first systematic reviews, Tao et al.¹³ analysed the clinical significance of ALHA preservation and recommended that close monitoring of postoperative liver function tests is essential, especially in patients with chronic liver disease or ALHA diameter >1.5 mm. They also found no significant difference in the number of retrieved lymph nodes between groups, similar to our results.

Limitations

The limitations of our study are that it was a single-center, retrospective design, and unlike some similar studies, the aberrant artery diameter was not studied. However, the strengths of our study include the inclusion of two major subgroups, total and subtotal gastrectomy, and the comparison of cases performed with open, laparoscopic and robotic techniques, unlike studies in the literature.

CONCLUSION

In our retrospective analysis, the prevalence of ALHA was 17.2% in patients with gastric cancer. Although we found a transient increase in postoperative liver enzymes in the ligated group, the statistical significance between the groups disappeared on day 30. We recommend close postoperative monitoring of liver enzymes in patients with suspected ALHA and collaboration with an experienced radiologist. Our results may contribute to the point of view that ALHA ligation is a safe intervention in gastric cancer surgery in terms of liver function.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Başakşehir Çam and Sakura City Hospital Clinical Researches Ethics Committee (Date: 22.11.2023, Decision No: 602).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394-424. doi:10.3322/caac.21492
- Rawla P, Barsouk A. Epidemiology of gastric cancer: global trends, risk factors and prevention. *Prz Gastroenterol*. 2019;14(1):26-38. doi:10.5114/pg.2018.80001
- Japanese Gastric Cancer Association. Japanese gastric cancer treatment guidelines 2018 (5th edition). *Gastric Cancer*. 2021;24(1):1-21. doi:10.1007/s10120-020-01042-y
- Sun YQ, Zhong Q, Lv CB, et al. The safety and efficacy of neoadjuvant immunochemotherapy following laparoscopic gastrectomy for gastric cancer: a multicentre real-world clinical study. *Int J Surg*. 2024;110(8):4830-4838. doi:10.1097/JS9.0000000000001468
- Choi TW, Chung JW, Kim HC, et al. Anatomic variations of the hepatic artery in 5625 patients. *Radiol Cardiothorac Imaging*. 2021;3(4):e210007. doi:10.1148/ryct.2021210007
- Sheta AA. Prevalence of anatomical variants in the branches of celiac and superior mesenteric arteries among Egyptians. *Anat Cell Biol*. 2024;57(3):353-362. doi:10.5115/acb.23.316
- Cirocchi R, D'Andrea V, Amato B, et al. Aberrant left hepatic arteries arising from left gastric arteries and their clinical importance. *Surgeon*. 2020;18(2):100-112. doi:10.1016/j.surge.2019.06.002
- Ang RRG, Lee HJ, Bae JS, et al. Safety of ligation of aberrant left hepatic artery originating from left gastric artery in laparoscopic gastrectomy for gastric cancer. *Sci Rep*. 2020;10(1):5856. doi:10.1038/s41598-020-62587-7
- Kim J, Kim SM, Seo JE, et al. Should an aberrant left hepatic artery arising from the left gastric artery be preserved during laparoscopic gastrectomy for early gastric cancer treatment? *J Gastric Cancer*. 2016;16(2):72-77. doi:10.5230/jgc.2016.16.2.72
- Shinohara T, Ohyama S, Muto T, Yanaga K, Yamaguchi T. The significance of the aberrant left hepatic artery arising from the left gastric artery at curative gastrectomy for gastric cancer. *Eur J Surg Oncol*. 2007;33(8):967-971. doi:10.1016/j.ejso.2007.02.030
- Huang CM, Chen QY, Lin JX, et al. Short-term clinical implications of the accessory left hepatic artery in patients undergoing radical gastrectomy for gastric cancer. *PLoS One*. 2013;8(5):e64300. doi:10.1371/journal.pone.0064300
- Oki E, Sakaguchi Y, Hiroshige S, Kusumoto T, Kakeji Y, Maehara Y. Preservation of an aberrant hepatic artery arising from the left gastric artery during laparoscopic gastrectomy for gastric cancer. *J Am Coll Surg*. 2011;212(5):e25-e27. doi:10.1016/j.jamcollsurg.2011.01.009
- Tao W, Peng D, Cheng YX, Zhang W. Clinical significance of aberrant left hepatic artery during gastrectomy: a systematic review. *World J Clin Cases*. 2022;10(10):3121-3130. doi:10.12998/wjcc.v10.i10.3121
- Lee S, Son T, Song JH, et al. Adverse effects of ligation of an aberrant left hepatic artery arising from the left gastric artery during radical gastrectomy for gastric cancer: a propensity score matching analysis. *J Gastric Cancer*. 2021;21(1):74-83. doi:10.5230/jgc.2021.21.e6
- Liu P, Wei M, Sun D, et al. Study on the application of preoperative three-dimensional CT angiography of perigastric arteries in laparoscopic radical gastrectomy. *Sci Rep*. 2022;12(1):6026. doi:10.1038/s41598-022-09584-0
- Sano A, Saito K, Kuriyama K, et al. Risk factors for postoperative liver enzyme elevation after laparoscopic gastrectomy for gastric cancer. *In Vivo*. 2021;35(2):1227-1234. doi:10.21873/invivo.12373
- Jeong GA, Cho GS, Shin EJ, Lee MS, Kim HC, Song OP. Liver function alterations after laparoscopy-assisted gastrectomy for gastric cancer and its clinical significance. *World J Gastroenterol*. 2011;17(3):372-378. doi:10.3748/wjg.v17.i3.372
- Ozutemiz C, Obuz F, Taylan A, Atila K, Bora S, Ellidokuz H. Volume change of segments II and III of the liver after gastrectomy in patients with gastric cancer. *Diagn Interv Radiol*. 2016;22(2):109-115. doi:10.5152/dir.2016.15143

Comparison of color stability of different 3D-printed resins

 Bengisu Koca,  Serdar Akarsu,  Sultan Aktuğ Karademir,  Samet Atasoy

Department of Restorative Dentistry, Faculty of Dentistry, Ordu University, Ordu, Türkiye

Cite this article as: Koca B, Akarsu S, Aktuğ Karademir S, Atasoy S. Comparison of color stability of different 3D-printed resins. *J Health Sci Med.* 2025;8(6):988-994.

Received: 31.07.2025

Accepted: 15.09.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to compare the long-term color stability of two different 3D-printed resins, one CAD/CAM resin nanoceramic block, and one nanohybrid composite resin after immersion in a coffee solution.

Methods: Specimens were prepared from four resin-based materials produced via different manufacturing techniques: subtractive manufacturing (Ceramart-group CS), additive manufacturing (VarseoSmile Crown Plus-group VS, Crowntec-group CT), and conventional methods (Clearfil Majesty Esthetic-group CME). The samples were immersed in coffee for 45 minutes daily over 30 days and stored in distilled water for the rest of the day. Color measurements were taken using a spectrophotometer on days 1, 7, 14, and 30. Color changes (ΔE) were calculated using the CIEDE2000 formula. Data were analyzed using one-way ANOVA, and post hoc comparisons were performed using the Tukey test. A p-value <0.05 was considered statistically significant.

Results: There were significant differences among the groups on days 1, 7 and 14. At the end of 30 days, the lowest color change was observed in group VS ($\Delta E=4.69$), while the highest was in group CS ($\Delta E=5.60$). The differences among groups were not statistically significant ($p=0.69$).

Conclusion: The color change of 3D-printed resins was comparable to that of CAD/CAM nanoceramic blocks and conventional composite resins. However, all materials exhibited color changes above the clinically acceptable threshold ($\Delta E>1.8$). Further studies are needed to evaluate the color stability of 3D printing resins and additive manufacturing techniques.

Keywords: 3D printing, CAD/CAM, composite resin, discoloration

INTRODUCTION

Composite resins are among the most widely used materials in restorative dentistry. Over time, improvements have been made to their esthetic properties, application techniques, and composition. However, discoloration over the long term remains a common clinical concern and a widely studied topic. Various factors can affect the color stability of restorative materials. Discoloration may occur via absorption or adsorption of staining substances from commonly consumed beverages such as coffee, tea, and wine. Therefore, the ability of restorative materials to resist color changes is crucial for the long-term clinical success of restorations.

Today, CAD/CAM manufacturing techniques can be classified into two main categories: subtractive and, more recently, additive manufacturing. Additive manufacturing, also known as 3D printing, eliminates the need for costly milling burs and blocks required in subtractive techniques, thereby significantly reducing post-production material waste. Unlike milling systems, dental 3D printers can produce multiple restorations simultaneously. Moreover, as 3D printing does not involve milling forces, it can be considered a passive production process.¹ The resins used in 3D printing are generally easier to repair and cause less wear on opposing dentition.² Some studies have also reported that restorations

produced via additive manufacturing exhibit greater accuracy compared to those made with subtractive techniques.³ Due to these advantages, 3D printing is increasingly used in various dental applications such as model fabrication, surgical guides, scaffold production, endodontics, orthodontics, removable prostheses, provisional crowns, inlays, onlays, overlays, full crowns, veneers, and implant planning.

Stereolithography (SLA) is the earliest and most widely used 3D printing technique in dentistry.^{1,3} SLA primarily enables the fabrication of resin-based objects. One of the most recent advancements in additive manufacturing is masked stereolithography (mSLA), which builds upon SLA principles. Unlike traditional SLA, which uses a laser as the light source, mSLA utilizes an array of LEDs that project UV light through a liquid crystal display (LCD), curing entire layers at once. Under certain conditions, mSLA can achieve faster print times than SLA due to the simultaneous curing of complete layers rather than individual points. Furthermore, the low cost of LCD screens makes mSLA a more affordable printing technology.⁴

The color stability of resin-based restorations in the oral environment depends not only on the manufacturing technique but also on exposure to dietary staining agents. Coffee is one

Corresponding Author: Serdar Akarsu, serdarakarsu@hotmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

of the most commonly consumed beverages worldwide due to its flavor and stimulating effects. Chromogens, which are pigments that adhere to enamel surfaces and cause visible stains over time, are the primary cause of coffee-related discoloration.⁵ Color mismatches in restorations are among the leading reasons for restoration replacements.

To overcome observer-related factors such as fatigue, ambient conditions, and subjective interpretation, modern research prefers spectrophotometers for color evaluation, offering more consistent and reproducible measurements. Since 2001, the CIEDE2000 formula, based on the CIELab* system, has been recommended to minimize the difference between perceived and calculated colors.⁶ The ΔE value represents the numerical difference between standardized L^* , a^* , and b^* values and the target color. It is commonly used to assess the acceptability and perceptibility thresholds of color differences between the restoration and the surrounding dentition.⁷

Although numerous studies have investigated the mechanical properties of 3D-printed resin materials, research focusing on discoloration—an essential aspect of esthetics—remains limited. Therefore, the aim of this in vitro study was to evaluate the color changes of a resin nanoceramic CAD/CAM block, a nanohybrid composite resin, and two different 3D-printed resins after daily immersion in a coffee solution. The null hypothesis was that the type of material would not significantly affect the degree of discoloration.

METHODS

Our study is a laboratory study. No patients or living tissues were used. Therefore, ethics committee approval is not required. All procedures were carried out in accordance with the ethical rules and the principles.

This in vitro study was conducted at the Laboratories of the Faculty of Dentistry, Ordu University, Department of Restorative Dentistry. Two different resins suitable for 3D printing with LCD technology, one CAD/CAM nanoceramic block, and a conventional nanohybrid composite were immersed in a coffee solution for 30 days. Color changes were evaluated on Days 1, 7, 14, and 30. The required sample size was determined using G*Power software version 3.1.9.2 (Heinrich Heine University, Düsseldorf, Germany) based on a 95% confidence level and a significance value of 0.05. The

sample size was determined to be 80 with 20 specimens for each group to account for potential specimen loss.

The workflow of the study included the following steps:

- Preparation of specimens
- Storage in distilled water and baseline color measurements
- Daily immersion in coffee solution and subsequent color measurements
- Comparison of color differences

The commercial names and compositions of the tested materials—VarseoSmile Crown Plus (BEGO, Bremen, Germany), Crowntec (Saremco, Switzerland), Cerasmart (GC Europe), and Clearfil Majesty Esthetic (Kuraray, Japan) are listed in [Table 1](#).

Preparation of Specimens

Group VS-VarseoSmile Crown Plus: The 3D-printed specimens were produced using the Phrozen Mini 8K printer (Phrozen Tech, Hsinchu, Taiwan), which employs high-resolution LCD technology. Disc-shaped specimens (8 mm diameter × 2 mm thickness) were designed using Blender software (Blender Foundation, Amsterdam, Netherlands) and exported in STL format. Files were processed in Chitubox v1.7.0 (Guangdong, China) with a 0° orientation and a 50 µm layer thickness. Six base layers were defined, with exposure parameters of 4.5 seconds per layer and 40 seconds for base exposure. Lift distance and speed were set to 8 mm and 60 mm/s, respectively. Support structures (Ø0.7 mm) were attached to a 0.8 mm thick curved base.

A total of 20 specimens were printed and cleaned using an ultrasonic cleaner; first in 95% reusable ethanol for 3 minutes, then in fresh 95% ethanol for 2 minutes. Residual resin was rinsed with ethanol spray and dried with compressed air. Support structures were removed using a separating disc (Dentorium, New York, USA). Post-curing was performed using a 405 nm UV light unit (Phrozen Cure, Phrozen 3D, Taiwan) for 2 × 5 minutes. Device specifications are listed in [Table 2](#).

Group CT-Crowntec: The same design, printer, and slicing parameters were used as in group VS. Exposure settings included 3.5–4.5 seconds per layer, 25–35 seconds for base layers, 6 mm lift distance, and 60 mm/min lift speed. The

Table 1. Resin-based restorative materials and their compositions used in the study

Manufacturing method	Material	Group name	Material type	Composition	Manufacturer	Lot. Num.
Additive manufacturing	VarseoSmile Crown Plus	VS	3D printed resin	4,4'-esterification products of isopropylidenediphenol, ethoxylated and 2-methylprop-2-enoic acid (50.00–<75.00 wt%), diphenyl (2,4,6-trimethylbenzoyl) phosphine oxide (<2.5 wt%)	VA; BEGO, Bremen, Germany	601003
	Crowntec	CT		Bis-EMA, trimethylbenzoyldiphenylphosphine oxide (0.1–1%)	Saremco Dental AG, Rebstein, Switzerland	E617
Subtractive manufacturing	Cerasmart	CS	Resin nanoceramic CAD/CAM Block	UDMA, Bis-EMA, silica (20 nm), barium oxide (0.1–1%)	GC dental products Europe, Leuven, Belgium	2104091
Conventional Fabrication	Clearfil Majesty Esthetic	CME	Nanohybrid conventional composite	Bisphenol A diglycidyl methacrylate (2.5–10%), silanized barium glass filler, prepolymerized organic filler, hydrophobic aromatic dimethacrylate, hydrophobic aliphatic dimethacrylate, DL-camphorquinone, accelerators, initiators, pigments	Kuraray, Japan	6U0250

VS: VarseoSmile, CT: Crowntec, CS: Cerasmart, CAD/CAM: Computer aided design/computer aided manufacturing, CME: Clearfil Majesty esthetic, EMA: Exponential moving average

Table 2. Additive manufacturing devices used in the study

Device type	Model	Technical specifications	Manufacturer
3D printer	Phrozen Sonic 8k Mini	Printing method: LCD Stereolithography Pixel size: 22 µm Pixel resolution: 7.536×3.240 Build volume (XYZ, mm): 165×72×180 Wavelength (nm): 405 Light Intensity (mW/cm ²): 2.8 Device weight (kg): 13	Phrozen Tech, Hsinchu, Taiwan
Washing unit	Phrozen Wash&Cure	Power output (W): 48 Input voltage (VDC): 24 Product dimensions (cm): 27.7×20.7×46.7 Washing tank capacity (L): 8 Max model size (cm): 21.8×12.3×23.5 Weight (kg): 3.5 Timer (min): 0-30	Phrozen 3D, Hsinchu City, Taiwan
Post-curing UV unit	Phrozen Wash&Cure	Power output (W): 48 Input voltage (VDC): 24 Product dimensions (cm): 35.4×30.9×36.8 Rotating plate diameter (cm): 19.6 Max model size (cm): 25×23.5 Weight (kg): 3.9 Timer (h): 0-2 LED wavelength (nm): 405	Phrozen 3D, Hsinchu City, Taiwan

LCD: Liquid crystal display

residual resin on the 20 specimens produced with these parameters was removed using a brush soaked in 95% ethanol until the surface achieved a matte appearance. Support structures were removed, and specimens were post-cured under the same UV lamp in two 5-minute cycles.

Group CS-CAD/CAM resin block (Cerasmart): Cylindrical blocks (8 mm diameter) were milled from Cerasmart blocks using a 5-axis CNC milling unit (Ceramill Motion 2, AmannGirrbach, Austria). The blocks were sectioned to 2 mm thickness using a precision diamond saw (Mecatome T180, France) under water cooling. 20 specimens were obtained.

Group CME-nanohybrid composite (Clearfil Majesty Esthetic): Composite discs were fabricated using silicone molds (8 mm diameter, 2 mm thick). Transparent strips were applied above and below the mold to ensure flat surfaces. The material was inserted using a spatula, compressed lightly, and polymerized with an LED curing unit according to the manufacturer's recommendations. Finishing and polishing were performed in two steps using the Clearfil Twist Dia system (Kuraray, Japan) for 15-20 seconds in a standardized direction.

Baseline Color Measurement

Baseline color measurements were obtained using a VITA EasyShade spectrophotometer (VITA Zahnfabrik, Bad Säckingen, Germany). The device was calibrated per manufacturer instructions before and between each measurement. Each sample was measured three times, and the average L*, a*, and b* values were recorded.

Aging Procedure in Coffee Solution

Each sample was immersed in a freshly prepared coffee solution (2 g Nescafe Classic in 200 ml boiling water, no sugar) for 45 minutes daily. The solution was cooled to room

temperature before immersion and renewed daily to prevent contamination.⁸ Specimens were stored in distilled water outside immersion periods and brushed under running tap water before each new cycle.

On days 1, 7, 14, and 30, specimens were brushed, dried, and color measurements were repeated. Measurements were made against an A4 white background with the device tip in direct perpendicular contact with the surface. The same observer conducted all measurements under identical room conditions and at the same time each day to ensure consistency.

Color Difference Calculation

Color differences were calculated using the CIEDE2000 formula:

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{K_L S_L}\right)^2 + \left(\frac{\Delta C'}{K_C S_C}\right)^2 + \left(\frac{\Delta H'}{K_H S_H}\right)^2 + R_T \left(\frac{\Delta C'}{K_C S_C}\right) \left(\frac{\Delta H'}{K_H S_H}\right)}$$

ΔL: Change in lightness

ΔC: Change in chroma

ΔH: Change in hue

SL, SC, SH: Weighting functions for perceptual uniformity

KL, KC, KH: Correction terms, each set to 1

RT: Rotation function accounting for chroma-hue interaction in blue areas

ΔE₀₀ value below 1.8 was considered the clinically acceptable threshold.⁹

Statistical Analysis

Data analyses were conducted using SPSS software (version 20.6, IBM Corp., Chicago, IL, USA) at a 95% confidence level. One-way ANOVA was used to compare groups, followed by Tukey's post hoc test for pairwise comparisons. Results were expressed as mean±standard deviation. Statistical significance was set at p<0.05.

RESULTS

On day 1, mean ΔE values were; group VS=1.10, group CT=1.53, group CS=1.81, and group CME=1.25. The differences between groups were statistically significant (p<0.0001). Group VS showed the lowest color change, and group CS the highest. There was no significant difference between group VS and group CME, whereas group VS and group CS differed significantly.

On day 7, ΔE values were: group VS=3.02, group CT=2.72, group CS=3.35, and group CME=2.47. Differences were again statistically significant (p<0.0001). The highest color change was observed in group CS, and the lowest in group CME. Significant differences were found between Crowntec and Cerasmart, and between VarseoSmile and Clearfil Majesty.

On day 14, ΔE values increased to; group VS=4.53, group CT=4.51, group CS=5.80, and group CME=3.78. The differences were statistically significant (p<0.0001). Group CME showed the lowest ΔE, and group CS the highest. Cerasmart differed significantly from all other groups, while the difference between 3D-printed and composite groups was not statistically significant.

On day 30, the mean ΔE values were, group VS=4.69, group CT=5.59, group CS=5.60, and group CME=5.20. No statistically significant differences were found among the

groups ($p=0.069$), although group VS still had the lowest and group CS the highest ΔE values

ΔE values for each group across all time points are summarized in Table 3 and illustrated in Figure.

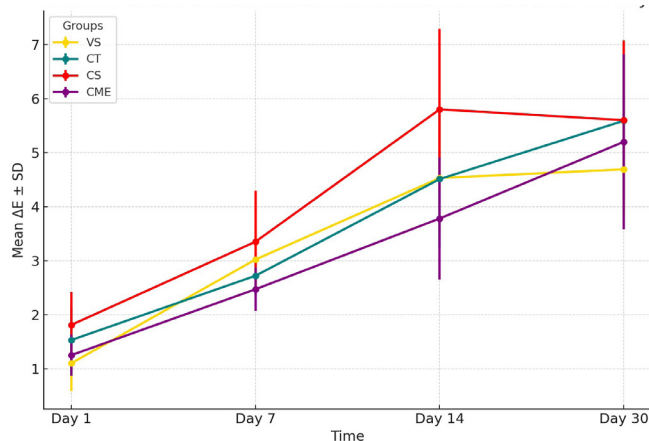


Figure. Mean ΔE values of resin-based restorative materials over a 30-day period

DISCUSSION

In this study, the optical properties of resin materials commonly used in additive manufacturing technology were investigated and compared with those of other resin-based restorative materials. After 30 days, no statistically significant differences were found among the groups, and the null hypothesis that the type of resin material does not affect color stability was accepted.

Resin materials with higher water absorption can also absorb greater amounts of water-soluble pigments.¹⁰ The color stability of resin materials has been improved through various physical and chemical modifications. These improvements include reduced water sorption, increased filler loading, smaller particle sizes, and an optimized filler-matrix interface.

In previous studies where CAD/CAM materials were immersed in distilled water, wine, and coffee for 30 days, Cerasmart samples stored in coffee showed ΔE values above the clinically acceptable threshold. These findings are consistent with the present study. Cerasmart's high translucency may explain the greater color change observed.^{11,12}

In resin-based composites, filler particle size and shape influence light scattering, while the presence of organic matrix and pigments affects light absorption.¹³ Studies have reported that the filler loading by weight of the 3D-printed resins Crowntec and VarseoSmile was 33%, within the range

of 30-50% as specified by the manufacturer.¹⁴ According to manufacturer data, their particle sizes are $0.7 \mu\text{m}$.² However, there is a lack of information on filler morphology. The detachment of inorganic particles from the surface over time may lead to void formation, increasing surface roughness. Additionally, the use of larger filler particles to reduce viscosity may cause sedimentation during long-term storage and heterogeneous layer printing during fabrication. A low filler content may increase light scattering and the volume of matrix requiring polymerization. Ideally, filler particles should be small and the filler content high to achieve optimal mechanical, physical, and optical properties. Smaller particles enhance light penetration.¹⁵ Cerasmart contains 71 vol% of silica (20 nm) and barium glass (300 nm) nanosized fillers with irregular shapes.¹⁶ Among CAD/CAM materials, Cerasmart is known to have a lower filler load. The excessive discoloration seen in the 30-day specimens may be attributed to this lower filler content. The lack of detailed chemical composition data on Crowntec and VarseoSmile limits the ability to correlate their optical behavior.

The monomers constituting the organic matrix also influence optical properties. In some studies, color stability was found to be more affected by the matrix than by filler content. According to limited manufacturer information, BisEMA-a monomer without hydroxyl groups-is used in the 3D printing resins. Another identified monomer, UDMA, has been shown to achieve three times greater polymerization than BisGMA.¹⁷ This reduces the amount of unpolymerized monomers, which are known to contribute to discoloration. The presence of BisEMA in all materials may explain the similar ΔE values observed in VarseoSmile, Crowntec, and Cerasmart.

All specimens were fabricated in A2 shade, as specified by the manufacturers. However, Alfaro et al.¹⁸ reported that many materials exhibit shades other than A2 when evaluated using the Vita Classic shade guide. This discrepancy may affect the reliability of color comparison studies.

Previous studies have shown that printer type can affect the color stability of resin-based restorations. In one study comparing surface roughness and stainability, Saremco Print Crowntec fabricated with DLP technology exhibited greater discoloration than Cerasmart.¹⁹ This study employed mSLA (LCD) printing technology. The primary distinction between mSLA and DLP/SLA methods lies in their curing mechanisms. mSLA uses an array of UV-LED or LED light sources masked by an LCD screen, directing light only to areas requiring curing. DLP projects light through movable mirrors, curing an entire layer simultaneously. These differences may explain why mSLA yields smoother and more detailed surfaces.²⁰

Table 3. Comparison of the mean ΔE values and standard deviations of four different resin materials over a 30-day period

	Group VS	Group CT	Group CS	Group CME	
	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	p-value*
Day 1	1.10 (0.51) ^A	1.53 (0.60) ^{AB}	1.81 (0.61) ^B	1.25 (0.38) ^A	<.0001
Day 7	3.02 (0.57) ^{AB}	2.72 (0.36) ^{BC}	3.35 (0.94) ^A	2.47 (0.40) ^C	<.0001
Day 14	4.53 (1.06) ^A	4.51 (1.26) ^A	5.80 (1.49) ^B	3.78 (1.13) ^A	<.0001
Day 30	4.69 (0.91) ^A	5.59 (0.53) ^A	5.60 (1.48) ^A	5.20 (1.62) ^A	.069

*One-way ANOVA results, groups with different magnitudes on each day were significantly different (Tukey HSD test). Uppercase letters within the rows indicate statistically significant differences between groups at the $p<0.05$ level. SD: Standard deviation

Ellakany et al.²¹ used a similar setup, comparing one subtractive CAD/CAM material, two additive SLA and DLP printed materials, and a conventional composite, all stored in various solutions for 1, 7, and 30 days. The DLP-printed specimens showed greater discoloration than SLA at all time points.

In another study, Krajangta et al.²² compared the color stability and surface roughness of 3D-printed VarseoSmile and subtractively fabricated Cerasmart stored in distilled water and coffee. The Asiga DLP printer was used. They reported no significant ΔE difference between the two after 30 days, but both exceeded the clinically acceptable limit.

Due to the layer-by-layer process of additive manufacturing, printed objects exhibit anisotropic mechanical properties.²³ Anisotropy refers to property variation depending on the printing orientation.²⁴ Print orientation affects the number of layers and therefore influences both mechanical and optical properties. A surface roughness greater than 0.2 μm may promote bacterial adhesion.²⁵ The layering lines may contribute to surface roughness depending on their location on the printed object. This may affect the roughness of the buccal surface, which is more exposed to staining agents.²³ Post-curing reduces anisotropy.²⁶

Tayaheri et al.²⁷ evaluated the degree of conversion in samples printed at 0° and 90°. Higher conversion rates were observed in the upper surfaces closer to the light source. Thus, 0° orientation places better-polymerized layers on the color measurement surface, while 90° does not, potentially affecting color stability. In this study, specimens were printed at 0°.

Post-curing has been shown to improve the mechanical and optical properties of 3D-printed resins. Lee et al.²⁸ investigated the effect of different post-curing durations (0, 5, 10, and 20 minutes) on color stability. As post-curing time increased, ΔE values decreased, particularly in samples immersed in coffee. Surface roughness also stabilized after 10 minutes of curing. In the present study, a 10-minute post-curing protocol was followed according to manufacturer instructions.

It is generally accepted that most indirect resin materials exhibit better color stability than direct resin-based materials due to their higher degree of conversion. However, in our study, no statistically significant difference was found among the ΔE values of all specimens.

Light-cured resin materials never achieve complete monomer conversion due to the gel effect, which reduces monomer mobility and limits polymerization. Therefore, post-curing is a critical step for additive resins.²⁹

Among all groups, group CS (Cerasmart) showed the highest discoloration at all time points. This may be attributed to milling marks remaining on the surface. Shin et al.²⁰ found that CAD/CAM materials retained milling lines, whereas 3D-printed samples exhibited smoother surfaces under high magnification. No polishing was applied at that stage, highlighting the importance of polishing protocols in additive manufacturing. Although 3D-printed materials may be expected to demonstrate superior color stability due to smoother surfaces, multiple factors-including the fabrication and post-processing steps-play a role.

Polishing systems vary in abrasive composition and particle characteristics, affecting their efficiency.³⁰ In this study, polishing was performed on all specimens using a two-step Twist Dia polishing kit (Clearfil Twist Dia, Kuraray, Japan) to ensure standardization among the groups

Avunduk et al.³¹ polished Crowntec and Clearfil Majesty Esthetic using one-step and multi-step procedures before immersing them in coffee. Although there were statistically significant differences in surface roughness, no significant ΔE differences were observed.

Özer et al.³² compared the ΔE values of 3D-printed (Crowntec, VarseoSmile) and CAD/CAM (Cerasmart) materials immersed in coffee, tea, and artificial saliva. On day 1, Cerasmart showed lower ΔE values than the 3D-printed resins, but on day 7, the values were 4.87 for Cerasmart, 4.70 for Crowntec, and 6.00 for VarseoSmile. In contrast to the present study, Sof-Lex discs were used for polishing, and samples were immersed continuously, rather than for 45 minutes per day. These factors likely contributed to the differing results.

An ideal resin material should be chemically stable and impermeable to water. Adverse effects include mechanical property degradation, volume changes, discoloration, reduced wear resistance, and hydrolysis at the filler-matrix interface. Prolonged immersion increases water absorption, accelerating color change in both conventional and CAD/CAM materials.³³

Coffee lowers pH due to its tannic acid and solvent content, increasing surface roughness.⁵ The pigments in coffee have low polarity and readily penetrate polymer matrices.⁶ These properties contribute to coffee's stronger staining effect compared to other beverages. Therefore, coffee was selected as the staining agent in this study.

Color measurements were made using a spectrophotometer. When properly calibrated before each measurement, the Vita EasyShade spectrophotometer offers high precision and reliability.³⁴ However, its calibration on opaque surfaces raises concerns about accuracy when measuring translucent resins.¹⁰

Nagai et al.³⁵ assessed discoloration in VarseoSmile and Cerasmart immersed in red wine for six months. VarseoSmile showed ΔE values above the clinical threshold, while Cerasmart did not. This discrepancy may be due to the use of a different device (Konica Minolta CM-2600d), which features an integrated light-diffusing sphere that reduces edge scattering.³⁴

The CIEDE2000 formula, which refines the traditional CIELab* method, offers improved accuracy in color difference calculation.⁷ This formula was used in the present study, with the clinical acceptability threshold set at $\Delta E < 1.8$.

In modern digital dentistry, 3D printers and printable resins are increasingly utilized, yet many questions remain regarding their behavior. Although recent studies have examined the discoloration of computer-aided resins, results vary significantly. This is likely due to the many variables affecting the mechanical, physical, and optical properties of 3D-printed dental materials. These include software, printer type, printing parameters, resin type, storage conditions, ambient

temperature, washing and post-curing processes, light source, curing duration, and the organic and inorganic composition of the resin. Further research is needed to establish reliable clinical protocols and indications.

Limitations

This study has certain limitations. This study only used a white background to measure colour stability. The prepared coffee solution cannot fully simulate the oral environment. The effects of saliva, smoking, and oral hygiene were not considered. Furthermore, while only the polished surface of restorations is typically exposed in vivo, both surfaces of the samples in this study were immersed, possibly leading to greater discoloration than would occur clinically.

CONCLUSION

Statistically significant differences were observed among the groups on days 1 and 7, where Cerasmart (CS) showed higher color change compared to the other materials. On day 14, intergroup differences persisted, while no significant difference was found on day 30 (>0.05). All materials showed color changes above the clinically acceptable threshold. Clinically, these findings suggest that the choice of material may influence early color stability, and careful selection should be considered during the first weeks after restoration. Further clinical studies are needed to investigate the effects of printer type, post-curing treatments, and polishing on the color stability of 3D printed resin restorations.

ETHICAL DECLARATIONS

Ethics Committee Approval

Our study is a laboratory study. No patients or living tissues were used. Therefore, ethics committee approval is not required.

Informed Consent

Because the study has no study with human and human participants, no written informed consent form was obtained.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Jockusch J, Özcan M. Additive manufacturing of dental polymers: an overview on processes, materials and applications. *Dent Mater J*. 2020; 39(3):345-354. doi:10.4012/dmj.2019-123
- Sasany R, Donmez MB, de Paula MS, et al. Stainability and translucency of potassium aluminum sulfate applied computer-aided design and computer-aided manufacturing materials after coffee thermocycling. *J Esthet Restor Dent*. 2024;36(3):477-483. doi:10.1111/jerd.13154
- Bae EJ, Jeong ID, Kim WC, Kim JH. A comparative study of additive and subtractive manufacturing for dental restorations. *J Prosthet Dent*. 2017; 118(2):187-193. doi:10.1016/j.prosdent.2016.11.004
- Lopes AJ, Lee IH, MacDonald E, Quintana R, Wicker R. Laser curing of silver-based conductive inks for in situ 3D structural electronics fabrication in stereolithography. *J Mater Process Technol*. 2014;214(9): 1935-1945. doi:10.1016/j.jmatprotec.2014.04.009
- Kim S, Chung SH, Kim RJY, Park YS. Investigating the role of chlorogenic acids and coffee type in coffee-induced teeth discoloration. *Acta Odontol Scand*. 2024;82(1):1-8. doi:10.1080/00016357.2023.2245880
- Peñate L, Mercade M, Arregui M, Roig M, Basilio J, Cedeño R. Color stability of CAD/CAM interim material for long-term fixed dental prostheses vs. conventional materials after immersion in different staining solutions. *J Compos Sci*. 2021;5(4):106. doi:10.3390/jcs5040106
- Paravina RD, Pérez MM, Ghinea R. Acceptability and perceptibility thresholds in dentistry: a comprehensive review of clinical and research applications. *J Esthet Restor Dent*. 2019;31(2):103-112. doi:10.1111/jerd.12465
- Azmy E, Al-Kholy MRZ, Gad MM, Al-Thobity AM, Emam AM, Helal MA. Influence of different beverages on the color stability of nanocomposite denture base materials. *Int J Dent*. 2021;2021:5861848. doi:10.1155/2021/5861848
- Paravina RD, Ghinea R, Herrera LJ, et al. Color difference thresholds in dentistry. *J Esthet Restor Dent*. 2015;27(Suppl 1):S1-S9. doi:10.1111/jerd.12149
- Arocha MA, Mayoral JR, Lefever D, Mercade M, Basilio J, Roig M. Color stability of siloranes versus methacrylate-based composites after immersion in staining solutions. *Clin Oral Investig*. 2013;17(16):1481-1487. doi:10.1007/s00784-012-0837-7
- Aydın N, Karaoğlu S, Oktay EA, Kılıçarslan MA. Investigating the color changes on resin-based CAD/CAM blocks. *J Esthet Restor Dent*. 2020;32(2):251-256. doi:10.1111/jerd.12561
- Awad D, Stawarczyk B, Liebermann A, Ilie N. Translucency of esthetic dental restorative CAD/CAM materials and composite resins with respect to thickness and surface roughness. *J Prosthet Dent*. 2015;113(6): 534-540. doi:10.1016/j.prosdent.2014.12.003
- Pecho OE, Ghinea R, do Amaral EAN, Cardona JC, Della Bona A, Pérez MM. Relevant optical properties for direct restorative materials. *Dent Mater*. 2016;32(5):e105-e112. doi:10.1016/j.dental.2016.02.008
- Mudhaffer S, Althagafi R, Haider J, Satterthwaite J, Silikas N. Effects of printing orientation and artificial ageing on martens hardness and indentation modulus of 3D printed restorative resin materials. *Dent Mater*. 2024;40(7):1003-1014. doi:10.1016/j.dental.2024.05.005
- Vichi A, Balestra D, Scotti N, Louca C, Paolone G. Translucency of CAD/CAM and 3D printable composite materials for permanent dental restorations. *Polymers (Basel)*. 2023;15(6):1443. doi:10.3390/polym15061443
- Turker I, Yuzbasioglu E. Optical properties of CAD/CAM interpenetrating phase composites-an overview. *Glass Ceram*. 2024; 81(5):217-224.
- Lemon MT, Jones MS, Stansbury JW. Hydrogen bonding interactions in methacrylate monomers and polymers. *J Biomed Mater Res A*. 2007; 83(3):734-746. doi:10.1002/jbm.a.31448
- Umaña Alfaro G. Efecto del post curado en las propiedades físicas y mecánicas de resinas dentales de impresión 3D para restauraciones definitivas [master's thesis]. Ciudad: Universidad; 2024.
- Bozoğulları HN, Temizci T. Evaluation of the color stability, stainability, and surface roughness of permanent composite-based milled and 3D printed CAD/CAM restorative materials after thermocycling. *Appl Sci*. 2023;13(21):11895. doi:10.3390/app132111895
- Shin JW, Kim JE, Choi YJ, et al. Evaluation of the color stability of 3D-printed crown and bridge materials against various sources of discoloration: an in vitro study. *Materials (Basel)*. 2020;13(23):5359. doi: 10.3390/ma13235359
- Ellakany P, Fouda SM, AlGhamdi MA, Aly NM. Comparison of the color stability and surface roughness of 3-unit provisional fixed partial dentures fabricated by milling, conventional and different 3D printing fabrication techniques. *J Dent*. 2023;131:104458. doi:10.1016/j.jdent.2023.104458

22. Krajangta N, Klaisiri A, Leelaponglit S, Intralawan N, Tiansuwan P, Pisethsalasai N. Color alteration of CAD/CAM 3D-printed, milled resin-ceramic hybrid material compared to enamel. *Dent Mater J*. 2024; 43(3):386-393. doi:10.4012/dmj.2023-275
23. Puebla K, Arcaute K, Quintana R, Wicker RB. Effects of environmental conditions, aging, and build orientations on the mechanical properties of ASTM type I specimens manufactured via stereolithography. *Rapid Prototyp J*. 2012;18(5):374-388. doi:10.1108/13552541211250373
24. Unkovskiy A, Bui PH, Schille C, Geis-Gerstorfer J, Huettig F, Spintzyk S. Objects build orientation, positioning, and curing influence dimensional accuracy and flexural properties of stereolithographically printed resin. *Dent Mater*. 2018;34(12):e324-e333. doi:10.1016/j.dental.2018.09.011
25. Yadav RD, Raisingani D, Jindal D, Mathur R. A comparative analysis of different finishing and polishing devices on nanofilled, microfilled, and hybrid composite: a scanning electron microscopy and profilometric study. *Int J Clin Pediatr Dent*. 2016;9(3):201-206. doi:10.5005/jp-journals-10005-1364
26. Monzón M, Ortega Z, Hernández A, Paz R, Ortega F. Anisotropy of photopolymer parts made by digital light processing. *Materials (Basel)*. 2017;10(1):64. doi:10.3390/ma10010064
27. Tahayeri A, Morgan M, Fugolin AP, et al. 3D printed versus conventionally cured provisional crown and bridge dental materials. *Dent Mater*. 2018;34(2):192-200. doi:10.1016/j.dental.2017.10.003
28. Lee EH, Ahn JS, Lim YJ, Kwon HB, Kim MJ. Effect of post-curing time on the color stability and related properties of a tooth-colored 3D-printed resin material. *J Mech Behav Biomed Mater*. 2022;126:104993. doi:10.1016/j.jmbbm.2021.104993
29. Wada J, Wada K, Gibreel M, et al. Effect of surface polishing on physical properties of an occlusal splint material for additive manufacturing under protection gas post-curing condition. *Polymers (Basel)*. 2023; 15(3):625. doi:10.3390/polym15030625
30. Devlukia S, Hammond L, Malik K. Is surface roughness of direct resin composite restorations material and polisher-dependent? A systematic review. *J Esthet Restor Dent*. 2023;35(6):947-967. doi:10.1111/jerd.13102
31. Erturk-Avunduk AT, Atılan-Yavuz S, Filiz H, Cengiz-Yanardag E. A comparative study of polishing systems on optical properties and surface roughness of additively manufactured and conventional resin based composites. *Sci Rep*. 2024;14(1):25658. doi:10.1038/s41598-024-77449-9
32. Özer NE, Şahin Z, Çöl CY, Arık Ö, Kılıçarslan MA. The effect of different staining solutions on the color stability of permanent indirect composite resins produced by additive and subtractive techniques. *Meandros Med Dent J*. 2024;25(3):262-278. doi:10.69601/meandrosmdj.1533026
33. Alshali RZ, Salim NA, Satterthwaite JD, Silikas N. Long-term sorption and solubility of bulk-fill and conventional resin-composites in water and artificial saliva. *J Dent*. 2015;43(12):1511-1518. doi:10.1016/j.jdent.2015.10.001
34. Johnston WM. Color measurement in dentistry. *J Dent*. 2009;37(Suppl 1): e2-e6. doi:10.1016/j.jdent.2009.03.011
35. Nagai T, Alfaraj A, Chu TMG, Yang CC, Lin WS. Color stability of CAD-CAM hybrid ceramic materials following immersion in artificial saliva and wine. *J Prosthodont*. 2024;33(3):279-285. doi:10.1111/jopr.13868

The impact of the COVID-19 pandemic on gastric cancer stage and nutritional status

 Melih Can Gül¹,  Emre Hafızoğlu²

¹Department of Gastroenterology, Afyonkarahisar State Hospital, Afyonkarahisar, Türkiye

²Department of Medical Oncology, Afyonkarahisar State Hospital, Afyonkarahisar, Türkiye

Cite this article as: Gül MC, Hafızoğlu E. The impact of the COVID-19 pandemic on gastric cancer stage and nutritional status. *J Health Sci Med.* 2025;8(6):995-999.

Received: 24.07.2025

Accepted: 18.09.2025

Published: 25.10.2025

ABSTRACT

Aims: The COVID-19 pandemic disrupted diagnostic services worldwide, particularly gastrointestinal (GI) endoscopy, raising concerns about delayed diagnoses and stage migration in gastric cancer. *Helicobacter pylori* (*H. pylori*), a key factor in gastric carcinogenesis, may also influence mucosal susceptibility to SARS-CoV-2. However, data on how these disruptions affected gastric cancer profiles remain limited. To compare clinical, endoscopic, and pathological features of patients undergoing upper GI endoscopy during and after the COVID-19 pandemic, and to evaluate the association between nutritional status and tumor burden in gastric cancer patients.

Methods: This retrospective cohort study included 478 patients who underwent diagnostic upper GI endoscopy due to dyspeptic or alarm symptoms at a tertiary care center in Türkiye. Patients were divided into two groups: pandemic (June 2021 to June 2022; n=216) and post-pandemic (May to October 2023; n=262). Demographics, symptoms, endoscopic and histopathological findings, and body-mass index (BMI) were analyzed. Statistical methods included t-tests, Chi-square tests, logistic regression, and Spearman correlation.

Results: Patients evaluated during the pandemic more frequently reported alarm symptoms and had significantly lower weight and BMI. *H. pylori* positivity (41.7% vs. 28.2%; p=0.03), peptic ulcer (16.7% vs. 8.4%; p=0.04), and gastric cancer (7.4% vs. 2.3%; p=0.02) were all more prevalent in the pandemic group. Among 22 patients with gastric cancer, those diagnosed during the pandemic exhibited more advanced disease (T3-T4: 75.0% vs. 33.3%; p=0.02), more frequent nodal involvement (62.5% vs. 16.7%; p=0.01), and BMI. Pandemic-period endoscopy remained an independent predictor of advanced-stage gastric cancer in multivariable analysis (OR=3.41; 95% CI: 1.31-8.88; p=0.008). Strong inverse correlations were found between BMI and tumor depth (p=-0.73), and between BMI and nodal involvement (p=-0.87).

Conclusion: The COVID-19 pandemic was associated with diagnostic delays and increased frequency of advanced-stage gastric cancer. BMI showed a strong inverse relationship with tumor aggressiveness and has potential to serve as a practical indicator of disease burden. These findings underscore the importance of preserving access to endoscopic diagnostics and highlight the prognostic value of nutritional parameters during public health emergencies.

Keywords: BMI, COVID-19, endoscopy, gastric cancer, *Helicobacter pylori*, nutritional status

INTRODUCTION

The COVID-19 pandemic has significantly disrupted global healthcare systems, leading to delays in diagnostic services, particularly in gastrointestinal (GI) endoscopy. During this period, access to upper GI endoscopy—a critical tool for diagnosing *Helicobacter pylori* (*H. pylori*), infection and gastric cancer—was substantially reduced due to infection control measures and institutional restrictions.¹⁻³ These delays have raised concerns about stage migration and missed opportunities for early detection of gastric malignancies.

H. pylori, the most prevalent chronic bacterial infection worldwide, plays a central role in the development of chronic gastritis, peptic ulcers, and gastric cancer. Emerging evidence also suggests a possible link between *H. pylori*-induced

intestinal metaplasia and increased susceptibility to SARS-CoV-2, due to upregulation of angiotensin-converting enzyme 2 (ACE-2) and Transmembrane Serine Protease 2 (TMPRSS-2) receptors in gastric mucosa.⁴⁻⁸ Although these mechanisms remain under investigation, the interplay between chronic mucosal inflammation and viral pathogenesis has gained relevance during the pandemic.

Gastric cancer remains a leading cause of cancer-related mortality globally, with prognosis largely dependent on stage at diagnosis.⁹ Prior studies have reported increased incidence of advanced-stage disease during the pandemic, likely reflecting diagnostic delays.^{10,11} Yet, real-world data comparing gastric cancer characteristics and nutritional

Corresponding Author: Melih Can Gül, opdrmelihcangul@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

status between pandemic and post-pandemic periods remain scarce, particularly from middle-income countries.

This study aimed to evaluate the clinical, endoscopic, and pathological features of patients undergoing upper GI endoscopy during and after the COVID-19 pandemic. Specifically, we investigated differences in symptom profiles, *H. pylori* prevalence, and gastric cancer stage, and assessed the relationship between nutritional status and tumor burden. By addressing this gap, our findings provide timely insight into the oncologic consequences of diagnostic disruptions and highlight potential prognostic value of simple clinical parameters in gastric cancer.

METHODS

Study Design and Ethics

This retrospective cohort study was conducted at the Endoscopy Unit of a tertiary referral hospital in Türkiye, which also served as a COVID-19 center during the pandemic. The study was approved by the Afyonkarahisar Health Sciences University Clinical Researches Ethics Committee (Date: 01.07.2022, Decision No: 2022/8). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki. Data were prospectively collected and retrospectively analyzed to compare outcomes between the pandemic and post-pandemic periods. The post-pandemic phase was defined according to the World Health Organization's declaration on 5 May 2023.

Study Population and Data Collection

This study included patients who underwent upper GI endoscopy due to dyspeptic or alarm symptoms between June 2021 and October 2023. Based on the timing of endoscopy, two groups were defined: the pandemic group, comprising patients examined between June 2021 and June 2022 (n=216), and the post-pandemic group, consisting of patients evaluated between May 2023 and October 2023 (n=262).

Eligible participants met the following inclusion criteria: age ≥ 18 years, presence of relevant upper GI symptoms, diagnostic upper GI endoscopy performed, and availability of complete clinical, endoscopic, and histopathological records. The exclusion criteria were: endoscopy performed for screening purposes, prior history of *H. pylori* treatment or gastric cancer, presence of any other active malignancy, incomplete data, and age < 18 years.

Demographic and clinical information-including age, sex, body weight, height, body-mass index (BMI), COVID-19 infection history, and vaccination status-was extracted from institutional electronic medical records. Presenting symptoms such as nausea, vomiting, hematemesis, retrosternal burning, bloating, melena, and weight loss were recorded from standardized intake forms.

Endoscopic findings included *H. pylori* infection, which was defined as a Sydney system score ≥ 2 according to the updated Sydney classification,¹² as well as peptic ulcers, gastric polyps, lower esophageal sphincter (LES) dysfunction, and any lesions suspicious for malignancy. Histopathological variables comprised tumor localization, histologic grade, T-stage, nodal involvement, and tumor, node, metastasis (TNM) staging

according to the 8th edition of the American Joint Committee on Cancer (AJCC) classification.

Statistical Analysis

All data analyses were conducted using IBM SPSS Statistics version 26.0 (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as mean $\pm$ standard deviation (SD), while categorical variables were presented as number and percentage. The normality of distribution for continuous variables was assessed using the Kolmogorov-Smirnov test and visual inspection of histograms.

For comparison between groups, the independent samples T test was applied to normally distributed continuous variables. Categorical variables were compared using the Chi-square test, or Fisher's exact test when expected cell counts were below 5. All tests were two-sided, and a p-value < 0.05 was considered statistically significant. Given the exploratory nature of the study, no corrections for multiple comparisons were applied.

To investigate the independent association between undergoing endoscopy during the pandemic and advanced-stage gastric cancer (defined as T3-T4), a binary logistic regression model was constructed. Covariates included in the model were the pandemic period (yes/no), age, sex, and BMI. These variables were selected based on clinical relevance and univariate p-values < 0.10 . Due to the limited number of gastric cancer cases (n=22), model complexity was restricted following the events-per-variable (EPV) rule, with a maximum of three predictors to avoid overfitting. Model adequacy was assessed using the Hosmer-Lemeshow goodness-of-fit test (p > 0.05), and multicollinearity was excluded by ensuring all variance inflation factor (VIF) values were < 2 . Adjusted odds ratios (ORs) with 95% confidence intervals (CIs) were reported.

Spearman's rank correlation coefficients (p) were calculated to assess monotonic associations among continuous and ordinal clinicopathological variables, including age, BMI, tumor stage, T-stage, and nodal involvement in gastric cancer patients. Correlation strength was interpreted as weak (p=0.10-0.39), moderate (p=0.40-0.69), or strong (p ≥ 0.70). Due to the exploratory nature and limited sample size, no causal inference was made from these correlations, and findings were interpreted with caution.

Missing data were handled using a complete-case analysis approach. Patients with missing values in any primary clinical, endoscopic, or histopathological variable were excluded prior to analysis. The overall missingness across variables was less than 5%, and no imputation techniques were employed.

Use of Generative Artificial Intelligence

OpenAI's ChatGPT was used solely for grammar correction and format editing. No AI tool contributed to data analysis, interpretation, or scientific decision-making. All results and conclusions were derived independently by the research team.

RESULTS

Between March 2020 and June 2021 (pandemic period) and July 2021 to December 2022 (post-pandemic period), a total

of 478 patients underwent upper gastrointestinal endoscopy at our center. Of these, 216 patients were evaluated during the pandemic and 262 in the post-pandemic period.

The pandemic group was older on average than the post-pandemic group (64.2 ± 10.3 vs. 59.8 ± 9.6 years; $p=0.051$). Mean body weight and BMI were significantly lower in the pandemic group (57.3 ± 8.4 kg and 22.4 ± 3.0 kg/m²) compared to the post-pandemic group (62.1 ± 9.2 kg and 24.0 ± 3.1 kg/m²; $p=0.047$ and $p=0.032$, respectively). No significant differences were found regarding height, sex, prior COVID-19 infection, or vaccination status (Table 1).

Table 1. Demographic characteristics and COVID-19-related variables of patients undergoing upper gastrointestinal endoscopy

Variable	Pandemic group (n=216)	Post-pandemic group (n=262)	p-value
Age (years)	64.2±10.3 (42-85)	59.8±9.6 (40-83)	0.051
Weight (kg)	57.3±8.4 (42-72)	62.1±9.2 (45-78)	0.047
Height (cm)	160.0±6.0 (150-170)	161.0±6.5 (152-172)	0.10
BMI (kg/m ²)	22.4±3.0 (17.1-28.3)	24.0±3.1 (18.2-29.4)	0.032
Female, n (%)	114 (52.8%)	138 (52.7%)	0.98
History of COVID-19	90 (41.7%)	98 (37.4%)	0.57
Vaccinated	76 (35.2%)	104 (39.7%)	0.58

Note: Values are presented as mean±standard deviation (range) or number (%), as appropriate. Statistical test: Independent samples T test or Chi-square test were used. BMI: Body-mass index

Upper gastrointestinal symptoms differed significantly between groups (Table 2). Nausea, vomiting, weight loss, melena, and hematemesis were more common during the pandemic period (all $p<0.01$), while retrosternal burning and bloating were more frequent post-pandemic ($p<0.01$ and $p=0.05$, respectively).

Table 2. Distribution of upper gastrointestinal symptoms in patients during and after the COVID-19 pandemic

Symptom	Pandemic group (n=216)	Post-pandemic group (n=262)	p-value
Nausea	118 (54.6%)	78 (29.8%)	<0.01
Vomiting	86 (39.8%)	52 (19.8%)	<0.01
Weight loss	130 (60.2%)	92 (35.1%)	<0.01
Melena	56 (25.9%)	18 (6.9%)	<0.01
Hematemesis	40 (18.5%)	10 (3.8%)	<0.01
Retrosternal burning	52 (24.1%)	112 (42.7%)	<0.01
Bloating	82 (38.0%)	132 (50.4%)	0.05

Note: Values are presented as number (percentage). Statistical test: Chi-square test was used

Endoscopic and histopathological findings showed higher rates of *H. pylori* infection (41.7% vs. 28.2%; $p=0.03$), peptic ulcer (16.7% vs. 8.4%; $p=0.04$), and gastric cancer (7.4% vs. 2.3%; $p=0.02$) during the pandemic (Table 3). Conversely, benign findings such as gastric polyps and LES dysfunction were more frequently detected after the pandemic ($p=0.03$ and $p=0.04$, respectively).

In total, 22 patients were diagnosed with gastric cancer (n=16 in the pandemic group and n=6 post-pandemic). While limited by small numbers, comparison revealed that early-

Table 3. Comparison of endoscopic and histopathological findings between pandemic and post-pandemic periods

Finding	Pandemic group (n=216)	Post-pandemic group (n=262)	p-value
<i>Helicobacter pylori</i>	90 (41.7%)	74 (28.2%)	0.03
Peptic ulcer	36 (16.7%)	22 (8.4%)	0.04
Gastric cancer	16 (7.4%)	6 (2.3%)	0.02
Gastric polyps	6 (2.8%)	18 (6.9%)	0.03
LES dysfunction	20 (9.3%)	40 (15.3%)	0.04

Note: Values are presented as number (percentage). Statistical test: Chi-square test was used. Definition: *Helicobacter pylori* positivity was defined as a Sydney score of ≥ 2

stage tumors (Stage I) were more often detected post-pandemic (33.3% vs. 12.5%; $p=0.01$), whereas advanced stages (Stage IV: 22.4% vs. 5.8%, $p=0.04$; T3-T4 depth: 75.0% vs. 33.3%, $p=0.02$) and nodal involvement (62.5% vs. 16.7%; $p=0.01$) were more frequent during the pandemic. Although more cases in the pandemic group were poorly differentiated (50.0% vs. 33.3%), this was not statistically significant ($p=0.22$) (Table 4).

Table 4. Clinicopathological features and staging of gastric cancer cases diagnosed during and after the COVID-19 pandemic

Variable	Pandemic group (n=16)	Post-pandemic group (n=6)	p-value
Stage I	2 (12.5%)	2 (33.3%)	0.01
Stage II	2 (12.5%)	1 (16.7%)	0.45
Stage III	8 (52.6%)	3 (44.2%)	0.08
Stage IV	4 (22.4%)	2 (5.8%)	0.04
T3-T4 tumors	12 (75.0%)	2 (33.3%)	0.02
Nodal involvement	10 (62.5%)	1 (16.7%)	0.01
Poor differentiation	8 (50.0%)	2 (33.3%)	0.22

Note: Values are presented as number (percentage). Statistical test: Chi-square test was used

To explore whether undergoing endoscopy during the pandemic was independently associated with advanced gastric cancer, a multivariable logistic regression analysis was performed. After adjusting for age, sex, and BMI, presentation during the pandemic period remained significantly associated with T3-T4 disease (OR=3.41; 95% CI: 1.31-8.88; $p=0.008$) (Figure). Due to the limited number of events, model overfitting was carefully avoided by including only essential covariates based on clinical relevance and univariable screening.

Cancer incidence according to pandemic status

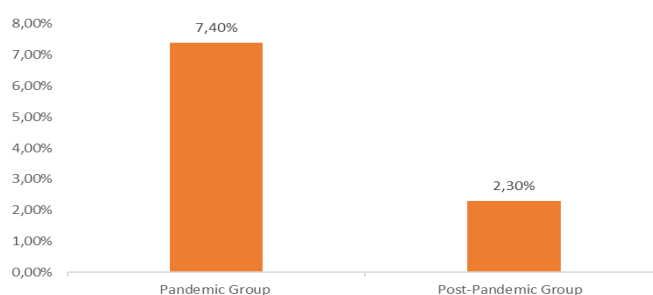


Figure. Pandemic period as an independent risk factor for advanced-stage gastric cancer

Multivariable logistic regression analysis showing the independent association between upper gastrointestinal endoscopy performed during the COVID-19 pandemic and advanced-stage gastric cancer (T3-T4). The model was adjusted for age, sex, and body mass index (BMI). Odds ratio (OR)=3.41; 95% confidence interval (CI): 1.31-8.88; $p=0.008$

Spearman's correlation analysis supported several inter-variable relationships (Table 5). Disease stage was positively correlated with tumor depth ($p=0.59$, $p=0.004$), nodal involvement ($p=0.60$, $p=0.003$), and age ($p=0.71$, $p=0.001$), and negatively with BMI ($p=-0.55$, $p=0.008$). T3-T4 invasion also correlated strongly with nodal involvement ($p=0.76$, $p<0.001$) and inversely with BMI ($p=-0.73$, $p=0.001$). While these associations do not imply causation, they highlight important patterns in the observed data. Notably, nodal involvement was strongly and inversely correlated with BMI ($p=-0.87$, $p=0.001$), and moderately associated with age ($p=0.47$, $p=0.029$).

DISCUSSION

This study is among the few from Türkiye to investigate the impact of the COVID-19 pandemic on the clinical and pathological presentation of gastric cancer. Patients evaluated during the pandemic were more likely to present with alarm symptoms, lower BMI, higher *H. pylori* prevalence, and more advanced tumors. Multivariable analysis demonstrated that undergoing endoscopy during the pandemic independently increased the risk of T3-T4 gastric cancer by more than threefold, even after adjustment for age, sex, and BMI. These findings support the hypothesis that pandemic-related diagnostic delays contributed to more advanced disease at diagnosis.

Patients in the pandemic group also presented with significantly lower BMI, deeper tumor invasion, and more frequent nodal involvement. Spearman's analysis confirmed strong inverse correlations between BMI and both T-stage and nodal spread, particularly for nodal involvement ($p=-0.87$, $p=0.001$). While causality cannot be established, these relationships may carry prognostic value and warrant further investigation in prospective studies.

Our results are consistent with previous reports indicating reduced access to endoscopy and increased incidence of advanced-stage gastric cancer during the pandemic.¹³⁻¹⁹ Professional societies such as European Association for Endoscopic Surgery recommended deferring elective endoscopies to reduce infection risk,^{13,14} a measure that likely delayed cancer diagnoses. Notably, a Chinese study reported an increase in malignancy detection despite reduced endoscopy volumes, likely due to prioritization of high-risk patients.¹⁵

A study analyzing 987 gastroscopies during the pandemic versus 3,245 in the previous year found significantly higher detection rates of peptic ulcers, esophageal cancer, and gastric cancer in the pandemic cohort.²⁰ A statistically significant rise in the detection of *H. pylori* infection, peptic ulcer disease, and gastric malignancies was documented during the pandemic period.

An elevated prevalence of *H. pylori* infection was also observed during the pandemic. Molecular studies suggest that *H. pylori*-associated intestinal metaplasia upregulates ACE2 and TMPRSS2 expression in gastric mucosa, potentially facilitating SARS-CoV-2 entry.^{21,22} Regional data from Türkiye, Iraq, and other countries support these findings, showing increased *H. pylori* positivity and mucosal inflammation during the pandemic.²³⁻²⁵

In contrast, gastroesophageal reflux disease (GERD) related symptoms such as retrosternal burning were more frequently reported post-pandemic. This may reflect changes in dietary habits, psychosocial stress, or delayed access to elective care. Prior studies have documented persistent GI symptoms—including GERD, dyspepsia, and irritable bowel syndrome (IBS)—among non-hospitalized COVID-19 survivors.²⁶⁻²⁸

Malnutrition and cachexia are highly prevalent in gastric cancer, especially at advanced stages; one meta-analysis reported that up to 87% of gastric cancer patients have malnutrition and about 65-85% of those with advanced disease develop cancer cachexia.²⁹ Advanced tumors can induce systemic inflammatory and catabolic processes that drive cancer cachexia, leading to progressive weight and muscle loss, hypoalbuminemia, and immunonutritional decline.^{30,31} These effects provide a plausible biological explanation for the observed inverse correlation between BMI and tumor aggressiveness in our cohort.

Limitations

This study has limitations. Its retrospective, single-center design and the relatively small number of gastric cancer cases limit generalizability. Although survival data were available, they were not analyzed due to short follow-up and variability in cause of death. Additionally, defining cohorts based on time periods may introduce residual confounding related to evolving healthcare access and policy changes. Despite these limitations, our findings highlight the oncologic consequences of diagnostic delays during public health emergencies. Moreover, the strong associations between BMI, tumor depth, and nodal involvement suggest that simple clinical metrics may provide useful prognostic information, particularly in resource-limited or crisis settings.

CONCLUSION

The COVID-19 pandemic was associated with a shift toward more advanced-stage gastric cancer at diagnosis, likely driven by diagnostic delays and reduced access to endoscopic services. Patients evaluated during the pandemic had lower BMI, more alarm symptoms, higher *H. pylori* prevalence, and more aggressive tumor features. These findings highlight the importance of preserving access to gastrointestinal diagnostics during public health crises. BMI has potential to serve as a practical indicator of tumor burden in gastric cancer. However, the limited number of gastric cancer cases in our study reduces the generalizability of these results, and larger multicenter studies are needed.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Afyonkarahisar Health Sciences University Clinical Researches Ethics Committee (Date: 01.07.2022, Decision No: 2022/8).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Sahu T, Mehta A, Rathe YK, et al. Current understanding of the impact of COVID-19 on gastrointestinal disease: challenges and openings. *World J Gastroenterol*. 2021;27(6):449-469. doi:10.3748/wjg.v27.i6.449
- Gralnek IM, Dumonceau JM, Kuipers EJ, et al. Diagnosis and management of nonvariceal upper gastrointestinal hemorrhage: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. *Endoscopy*. 2015;47(10):a1-a46. doi:10.1055/s-0034-1393172
- Han X, Yang NN, Nogueira L, et al. Changes in cancer diagnoses and stage distribution during the first year of the COVID-19 pandemic in the USA: a cross-sectional nationwide assessment. *Lancet Oncol*. 2023;24(8):855-867. doi:10.1016/S1470-2045(23)00293-0
- Sipponen P, Hyvärinen H. Role of *Helicobacter pylori* in the pathogenesis of gastritis, peptic ulcer and gastric cancer. *Scand J Gastroenterol Suppl*. 1993;196:3-6. doi:10.3109/00365529309098333
- Ansari S, Yamaoka Y. *Helicobacter pylori* virulence factor cytotoxin-associated gene A (CagA)-mediated gastric pathogenicity. *Int J Mol Sci*. 2020;21(19):7430. doi:10.3390/ijms21197430
- Bhattacharjee A, Sahoo OS, Sarkar A, et al. Infiltration to infection: key virulence players of *Helicobacter pylori* pathogenicity. *Infection*. 2024;52(2):345-384. doi:10.1007/s15010-023-02159-9
- Tandon A, Baral B, Saini V, et al. The role of *Helicobacter pylori* in augmenting the severity of SARS-CoV-2 related gastrointestinal symptoms: an insight from molecular mechanism of co-infection. *Heliyon*. 2024;10(18):e37585. doi:10.1016/j.heliyon.2024.e37585
- Durazzo M, Adriani A, Fagoonee S, Saracco GM, Pellicano R. *Helicobacter pylori* and respiratory diseases: 2021 update. *Microorganisms*. 2021;9(10):2033. doi:10.3390/microorganisms9102033
- Shirani M, Pakzad R, Haddadi MH, et al. The global prevalence of gastric cancer in *Helicobacter pylori*-infected individuals: a systematic review and meta-analysis. *BMC Infect Dis*. 2023;23(1):543. doi:10.1186/s12879-023-08504-5
- Qu HT, Li Q, Hao L, et al. Esophageal cancer screening, early detection and treatment: current insights and future directions. *World J Gastrointest Oncol*. 2024;16(4):1180-1191. doi:10.4251/wjgo.v16.i4.1180
- Shah R, Hanna NM, Loo CE, et al. The global impact of the COVID-19 pandemic on delays and disruptions in cancer care services: a systematic review and meta-analysis. *Nat Cancer*. 2025;6(1):194-204. doi:10.1038/s43018-024-00880-4
- Toyoshima O, Nishizawa T, Yoshida S, et al. Consistency between the endoscopic Kyoto classification and pathological updated Sydney system for gastritis: a cross-sectional study. *J Gastroenterol Hepatol*. 2022;37(2):291-300. doi:10.1111/jgh.15693
- Sethi A, Swaminath A, Latorre M, et al. Donning a new approach to the practice of gastroenterology: perspectives from the COVID-19 pandemic epicenter. *Clin Gastroenterol Hepatol*. 2020;18(8):1673-1681. doi:10.1016/j.cgh.2020.04.032
- Gralnek IM, Hassan C, Beilenhoff U, et al. ESGE and ESGENA position statement on gastrointestinal endoscopy and the COVID-19 pandemic. *Endoscopy*. 2020;52(6):483-490. doi:10.1055/a-1155-6229
- Zhu L, Cai MY, Shi Q, et al. Analysis of selective endoscopy results during the epidemic of Coronavirus disease 2019 (COVID-19). *Zhonghua Wei Chang Wai Ke Za Zhi*. 2020;23(4):327-331. doi:10.3760/cma.j.issn.1671-0274.2020-0316-00147
- Ma K, Chen X, Xiang X, et al. Willingness to undergo gastroscopy for early gastric cancer screening and its associated factors during the COVID-19 pandemic-a nationwide cross-sectional study in China. *Patient Prefer Adherence*. 2023;17:505-516. doi:10.2147/PPA.S400908
- Abu-Freha N, Hizkiya R, Abu-Abed M, et al. The impact of the COVID-19 pandemic on colorectal and gastric cancer diagnosis, disease stage and mortality. *Front Med (Lausanne)*. 2022;9:954878. doi:10.3389/fmed.2022.954878
- Fujita S, Sakuramoto S, Miyawaki Y, et al. Impact of the first era of the coronavirus disease 2019 pandemic on gastric cancer patients: a single-institutional analysis in Japan. *Int J Clin Oncol*. 2022;27(5):930-939. doi:10.1007/s10147-022-02142-8
- Arneiro AJ, Pereira MA, Dias AR, et al. Impact of COVID-19 pandemic on the surgical treatment of gastric cancer: a 3-year analysis. *Arq Bras Cir Dig*. 2025;37:e1868. doi:10.1590/0102-6720202400074e1868
- Huang K, Zhao X, Chen X, Gao Y, Yu J, Wu L. Analysis of digestive endoscopic results during COVID-19. *J Transl Intern Med*. 2021;9(1):38-42. doi:10.2478/jtim-2021-0006
- Zhang M, Feng C, Zhang X, et al. Susceptibility factors of stomach for SARS-CoV-2 and treatment implication of mucosal protective agent in COVID-19. *Front Med (Lausanne)*. 2020;7:597967. doi:10.3389/fmed.2020.597967
- Gonzalez I, Lindner C, Schneider I, et al. Inflammation at the crossroads of *Helicobacter pylori* and COVID-19. *Future Microbiol*. 2022;17(2):77-80. doi:10.2217/fmb-2021-0250
- Jasim Abdullah Y, Hasan NF, Zghair Jaber Alsaedi R. Correlation between *Helicobacter pylori* infection and COVID-19. *Iran J War Public Health*. 2021;13:255-259.
- Bozdağ E, Gülmez S, Yücesoy FS, et al. Analysis of upper gastrointestinal endoscopy results during the COVID-19 pandemic. *Ann Ital Chir*. 2022;93:391-397.
- Atak T, Erdoğan KO, Atak İ. An evaluation of gastroscopy findings during the COVID-19 pandemic: a public hospital-based study from Turkey. *Cureus*. 2023;15(10):e46451. doi:10.7759/cureus.46451
- Al-Momani H, Al Balawi D, Almasri M, et al. Gastroesophageal reflux in lockdown. *Future Sci OA*. 2023;9(6):FSO863. doi:10.2144/fsoa-2023-0042
- Ma Y, Zhang L, Wei R, et al. Risks of digestive diseases in long COVID: evidence from a population-based cohort study. *BMC Med*. 2024;22(1):14. doi:10.1186/s12916-023-03236-4
- Xu E, Xie Y, Al-Aly Z. Long-term gastrointestinal outcomes of COVID-19. *Nat Commun*. 2023;14(1):983. doi:10.1038/s41467-023-36223-7
- Hou S, Song D, Hao R, Li L, Zhang Y, Zhu J. Prognostic relevance of prognostic nutritional indices in gastric or gastro-esophageal junction cancer patients receiving immune checkpoint inhibitors: a systematic review and meta-analysis. *Front Immunol*. 2024;15:1382417. doi:10.3389/fimmu.2024.1382417
- Gül MC, Çolakoglu MK, Öter V, et al. Low preoperative cachexia index is associated with severe postoperative morbidity in patients undergoing gastrectomy for gastric cancer. *Diagnostics*. 2025;15(18):2284. doi:10.3390/diagnostics15182284
- Xu X, Tian M, Ding CC, Xu H, Wang H, Jin X. Skeletal Muscle Index-Based Cachexia Index as a predictor of prognosis in patients with cancer: a meta-analysis and systematic review. *Nutr Rev*. 2025;83(3):e852-e865. doi:10.1093/nutrit/nuae094

The mediating role of perceived stress: an examination of the relationship between health literacy and healthy living awareness

✉Müjdat Yeşildal¹, ✉Adil Aydoğdu¹, ✉Rumeysa Çelen², ✉Tunahan Yaver²

¹Department of Health Management, Faculty of Health Sciences, Selçuk University, Konya, Türkiye

²Master's Student, Institute of Health Sciences, Selçuk University, Konya, Türkiye

Cite this article as: Yeşildal M, Aydoğdu A, Çelen R, Yaver T. The mediating role of perceived stress: an examination of the relationship between health literacy and healthy living awareness. *J Health Sci Med.* 2025;8(6):1000-1005.

Received: 04.08.2025

Accepted: 18.09.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to explore the mediating effect of perceived stress on the relationship between health literacy and healthy lifestyle awareness. It was hypothesized that health literacy would positively predict healthy lifestyle awareness, directly and indirectly, by reducing levels of perceived stress.

Methods: This research adopted a cross-sectional and correlational design. The sample consisted of 1,150 adults residing in Konya, Türkiye. Data collection instruments included validated measures of Health Literacy, Perceived Stress, and Healthy Life Awareness. Descriptive analyses and Pearson correlation coefficients were calculated, followed by mediation testing through structural equation modeling. The significance of indirect effects was assessed using bootstrapping with 5,000 resamples and 95% confidence intervals.

Results: The sample consisted of 1,150 adults residing in Konya, with a mean age of 30.48. The majority were female (54.3%), single (59.7%), and university graduates (59.1%). The findings demonstrated that health literacy positively predicted healthy lifestyle awareness ($\beta=0.43$, $p<0.001$) and negatively predicted perceived stress ($\beta=-0.21$, $p<0.001$). In turn, perceived stress negatively predicted healthy lifestyle awareness ($\beta=-0.19$, $p<0.001$). Mediation analysis revealed a significant indirect effect of health literacy on healthy lifestyle awareness through perceived stress [$\beta=0.02$, 95% CI (0.01-0.04)]. These results supported the hypothesized mediating model.

Conclusion: The study proves that perceived stress partially mediates the relationship between health literacy and healthy lifestyle awareness. Improving health literacy may reduce perceived stress, enhancing individuals' understanding of healthy living. These findings highlight the importance of incorporating stress management and health literacy education into university health programs to promote holistic well-being.

Keywords: Health literacy, stress, awareness, health behavior

INTRODUCTION

Individuals' general well-being and quality of life are closely related to their healthy living awareness (HLA). HLA encompasses the combined knowledge, attitudes, and behaviors that empower individuals to make informed health decisions, prevent illness, and lead a more mindful lifestyle.¹ A strong sense of health identity is an important motivator for individuals to maintain healthy behaviors.² In this regard, strengthening health education programs can contribute to positive changes in lifestyle choices by increasing individuals' knowledge levels.³

Health literacy (HL), the capacity to access, comprehend, and utilize health-related information, is considered one of the key cognitive factors shaping HLA.^{4,5} Extensive research has shown that individuals with higher HL levels are more effective in utilizing healthcare services, following treatment

recommendations, and attaining improved health outcomes.^{6,7} However, insufficient HL can lead to adverse outcomes such as exposure to misinformation, adoption of unhealthy lifestyles, and low utilization of health services.⁸ In addition, HL facilitates individuals' access to reliable information and protects them from misinformation.⁹ Various studies indicate that HL is positively associated with behaviors that promote health and well-being.¹⁰⁻¹²

However, psychological barriers may exist when translating health knowledge into behavior. In this context, perceived stress (PS) plays an important role. PS describes the mental and emotional reactions triggered when a person interprets a situation as potentially threatening or overwhelming.¹³ High stress levels can negatively affect individuals' mental functions and health awareness, preventing them from maintaining

Corresponding Author: Müjdat Yeşildal, mujdatyesildal@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

healthy lifestyle behaviors.^{14,15} The literature shows that stress determines individuals' cognitive and behavioral processes related to health.^{16,17} This situation shows that even individuals with high HL may find it challenging to translate this information into healthy lifestyle behaviors when under stress.¹⁸

The theoretical basis of this study is the Social Cognitive Theory, which explains individuals' behaviors through the interaction of environmental, cognitive, and personal factors, and the Stress-Perception Theory, which argues that stress is managed through individual perception and coping processes.^{13,19} In this context, testing the mediating role of PS in the relationship between HL and HLA provides a theoretically meaningful approach.

HL has been extensively studied concerning health behaviors and outcomes across various populations. Prior evidence demonstrates that higher HL is associated with better health decisions, treatment adherence, and preventive behaviors. However, most existing studies originate from Western or Asian contexts, and evidence from Türkiye remains limited. Moreover, while previous research has often established direct links between HL and healthy lifestyle behaviors, fewer studies have examined the psychological mechanisms underlying this relationship. In particular, the mediating role of PS in translating health knowledge into actual awareness and behavior has not been sufficiently explored. This gap is especially critical in Türkiye, where cultural dynamics, rapid digitalization of health information, and variations in health literacy levels may uniquely influence individuals' awareness of healthy living. By addressing this gap, the present study provides novel insights into how HL contributes to healthy lifestyle awareness through the psychological pathway of PS, thereby offering context-specific evidence for both academia and public health policy.

A limited body of research in Türkiye investigates the associations among HL, PS, and HLA. This study seeks to contribute to the literature by exploring how HL influences individuals' awareness of HLA, focusing on the mediating role of PS in this relationship. Structural equation modeling (SEM) was employed to test the hypothesized pathways. By clarifying the interaction among these variables, the study aims to enhance understanding of health-related decision-making processes and provide evidence-based insights for professionals and policymakers in shaping health education, stress reduction programs, and public health strategies.

METHODS

Ethical Approval

Ethical approval for this research was granted by the Selçuk University Faculty of Health Sciences Non-interventional Clinical Researches Ethics Committee (Date: 29.01.2025, Decision No: 2025/50). The study was conducted per the ethical principles of the Declaration of Helsinki. Before participating, individuals were fully informed about the purpose of the study, the voluntary nature of their involvement, and how their data would be kept confidential. Informed consent was obtained electronically, and participants were allowed

access to the questionnaire only after providing their explicit agreement to participate.

Research Design

This study used a cross-sectional and correlational research design to investigate whether PS mediates the relationship between HL and HLA. SEM was used to analyze the proposed relationships. The conceptual model, based on theoretical frameworks related to health behavior and stress processing, as shown in Figure, presents HL (X) as the independent variable, PS (M) as the mediating variable, and HLA (Y) as the dependent variable. Within this framework, the following hypotheses were tested:

- **H1:** HL significantly and positively predicts HLA.
- **H2:** HL significantly predicts PS in a negative direction.
- **H3:** PS significantly predicts HLA in a negative direction.
- **H4:** PS mediates the relationship between HL and HLA.

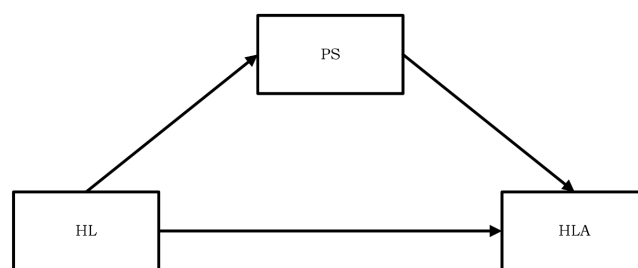


Figure. Research model

Participants and Sample

The study sample comprised individuals aged 18 and above living in the Konya province of Türkiye. 1,150 individuals were selected through convenience sampling between March and July 2025 and participated in the study. This sample size meets and exceeds the methodological recommendations commonly accepted for mediation analysis using bootstrapping techniques. In particular, mediation models involving multiple variables require a minimum of approximately 500 participants to obtain stable and robust estimates of indirect effects. Therefore, the current sample size is considered statistically sufficient regarding power and sensitivity to detect mediation effects.^{20,21}

Data Collection Tools

This study collected data using a structured questionnaire consisting of standardized and validated scales. The questionnaire included demographic questions and three main tools: the Health Literacy Scale, the Perceived Stress Scale, and the Healthy Living Awareness Scale. These tools were selected based on their relevance to the study variables and their psychometric properties established in previous research.

Personal Information Form

The researchers created a Personal Information Form for this study to gather data on participants' sociodemographic attributes. This form included categorical variables such as

gender, marital status, and chronic disease status (presence or absence of chronic disease). Additionally, continuous variables such as age (in years) and monthly income (in TRY) were recorded. These variables were included to characterize the sample and to examine potential covariates in statistical analyses.

Health Literacy Scale

This study used the Health Literacy Scale to measure individuals' perceived abilities to access, understand, evaluate, and apply health-related information. The scale consists of six items and has a four-point Likert-type structure. High scores on the scale indicate a high level of HL.^{22,23}

Perceived Stress Scale

The Perceived Stress Scale was used in this study to assess individuals' perceptions of stress. The scale consists of four items and is answered using a five-point Likert scale, with high scores indicating a high level of PS.^{24,25}

Healthy Life Awareness Scale

In this study, the Healthy Living Awareness Scale was used to measure individuals' levels of awareness regarding healthy living behaviors. The scale consists of 15 items and is answered using a five-point Likert Scale. High scores on the scale indicate that the individual has a higher level of awareness regarding healthy living issues.²⁶

Statistical Analysis

All data procedures were conducted using IBM SPSS Statistics version 30.0 and the PROCESS Macro 5.0 developed by Hayes (2022).²¹ Before hypothesis testing, the dataset was assessed for normality, multicollinearity, and bivariate associations among variables. Univariate normality was evaluated through skewness, kurtosis values, and the Kolmogorov-Smirnov test; values falling within the -1 to +1 range were considered indicative of approximate normal distribution.^{28,29} The internal consistency of the instruments was determined using Cronbach's alpha and McDonald's omega. Multicollinearity was examined using the Variance Inflation Factor (VIF) scores, with values below 5 suggesting no primary concern. Descriptive statistics and Pearson correlation coefficients were computed to explore the relationships among HL, PS, and HLA. The internal consistency of the instruments was determined using Cronbach's alpha, with α values of .70 or higher deemed acceptable. To address the main aim of the study, a simple mediation model (model 4) within the PROCESS Macro was employed, specifying HL as the independent variable (X), HLA as the outcome variable (Y), and PS as the mediator (M). Both direct and indirect effects were estimated using bias-corrected 95% confidence intervals derived from 5,000 bootstrap samples. Mediation was considered statistically significant if the confidence interval for the indirect effect did not include zero. A p-value of less than 0.05 was accepted as the criterion for statistical significance.

RESULTS

This section presents correlation analyses revealing the relationships between the descriptive statistics obtained in the

study and the variables and findings related to the mediation model.

Table 1 outlines the demographic profile of the 1,150 individuals who participated in the study. Among them, 54.3% were female and 45.7% were male. Most participants were single (59.7%), while 40.3% were married. In terms of educational attainment, the majority held a university degree (59.1%), followed by primary school (16.5%), high school (15.8%), and postgraduate qualifications (8.5%). 11.1% reported having at least one chronic disease, whereas 88.9% reported no chronic health conditions. The mean age of the participants was 30.48 years $\pm$ 12.34, with ages ranging from 18 to 78. The average monthly income was 45,616.21 $\pm$ 30,880.57, with reported incomes ranging between 14,500 and 165,000 TRY.

Table 1. Descriptive characteristics of the participants (n=1150)

Variables		n	%
Gender	Male	526	45.7
	Female	624	54.3
Marital status	Single	686	59.7
	Married	464	40.3
Educational status	Primary school	190	16.5
	High school	182	15.8
	Bachelor's degree	680	59.1
	Postgraduate	98	8.5
Chronic disease status	Yes	128	11.1
	No	1022	88.9
Variables	M (SD)	Min	Max
Age (years)	30.48 (12.34)	18	78
Monthly income (TRY)	45616.21 (30880.57)	14500	165000
SD: Standard deviation			

Table 2 summarizes the descriptive findings for the study variables, including mean scores, standard deviations, reliability estimates (Cronbach's α and McDonald's ω), and the correlations between the measured constructs. The mean score for HL was 2.89 $\pm$ 0.49, for PS it was 3.02 $\pm$ 0.64, and for HLA it was 3.77 $\pm$ 0.55. The reliability levels of the instruments were found to be acceptable, with α values ranging from 0.70 to 0.86 and ω values from 0.72 to 0.87, which are considered acceptable and indicative of good internal consistency according to conventional criteria. In terms of correlations, HL was positively and moderately correlated with HLA ($r=0.43$, $p<0.001$), indicating that individuals with higher HL tend to have greater awareness of healthy living. In contrast, HL was negatively correlated with PS ($r=-0.16$, $p<0.001$), suggesting that higher HL is associated with lower stress levels. Furthermore, PS demonstrated a negative correlation with HLA ($r=-0.19$, $p<0.001$), meaning increased stress is linked to lower awareness of healthy lifestyle behaviors. These correlation coefficients fall within the small-to-moderate range, providing preliminary evidence consistent with the hypothesized mediation model.

Table 2. Descriptive statistics, internal consistency coefficients, and correlations

	M (SD)	α	ω	HL	PS	HLA
HL	2.89 (0.49)	0.70	0.72	-		
PS	3.02 (0.64)	0.86	0.87	-0.16	-	
HLA	3.77 (0.55)	0.73	0.74	0.43	-0.19	-

SD: Standard deviation, HL: Health literacy, PS: Perceived stress, HLA: Healthy living awareness

Table 3 presents the mediation analysis results examining whether PS mediates the relationship between HL and HLA. The analysis revealed that HL had a significant negative effect on PS ($\beta=-0.16$, $SE=0.04$, $z=-5.60$, $p<0.001$; 95% CI=-0.29, -0.14), indicating that individuals with higher HL levels tend to experience lower levels of perceived stress. Likewise, PS was found to negatively and significantly influence HLA ($\beta=-0.12$, $SE=0.02$, $z=-4.61$, $p<0.001$; 95% CI=-0.15, -0.06), suggesting that increased stress is associated with reduced awareness of healthy lifestyle behaviors. The mediation pathway was also statistically significant, with PS partially mediating the link between HL and HLA ($\beta=0.02$, $SE=0.01$, $z=3.56$, $p<0.001$; 95% CI=0.01, 0.04). This implies that HL indirectly affects HLA through reductions in PS. Furthermore, HL's direct effect on HLA remained significant ($\beta=0.41$, $SE=0.03$, $z=15.23$, $p<0.001$; 95% CI=0.40, 0.52). The total effect of HL on HLA was estimated at $\beta=0.43$ ($z=16.05$, $p<0.001$; 95% CI=0.42, 0.54), affirming that HL is a strong and positive predictor of HLA. These results confirm that PS is a partial mediator in the relationship between HL and HLA.

All four hypotheses (H1-H4) examined in this study were statistically significant. The results revealed that HL is a positive predictor of wellness awareness (H1), HL is negatively associated with PS (H2), PS negatively predicts wellness awareness (H3), and PS partially mediates the relationship between HL and wellness awareness (H4). These outcomes are in harmony with the study's theoretical framework and provide empirical support for the proposed model.

DISCUSSION

The study's findings indicate that HL levels significantly predict HLA. According to the results, individuals' awareness of healthy living behaviors increases as HL increases. This finding reveals that individuals' capacity to acquire, evaluate, and integrate health-related information into their daily lives plays a decisive role in adopting a healthy lifestyle. Comparable results reported in the literature suggest that individuals with greater health literacy tend to make more

informed choices regarding their health and are more likely to engage in healthier lifestyle practices.^{4,10,11,27} In addition, it is emphasized that as HL levels increase, individuals tend to engage in health-protective and health-promoting behaviors. Nonetheless, certain studies indicate that this relationship may not be entirely direct or robust, and variables such as self-efficacy could function as mediators or moderators within this dynamic.²⁸ While HL plays a significant role in this context, it represents just one of several complex and interrelated factors that influence individuals' health-related behaviors.

Another important finding of the study is that HL levels significantly and negatively affect PS. It has been observed that individuals with easier access to health-related information and who can interpret this information correctly experience less stress when faced with health problems. This situation can be interpreted as the cognitive competencies gained through HL supporting self-efficacy in uncertainty and anxiety, thereby reducing stress levels.¹⁹ This finding is consistent with studies conducted on young adults and students, which report that depression, anxiety, and stress levels decrease as HL levels increase.^{15,29} Similarly, it has been reported that lower stress levels are observed in individuals with high levels of HL.³⁰ However, it is stated that the relationship between HL and stress may be affected by certain contextual variables. It is proposed that this association could differ based on variables such as the intensity of the individual's illness, the extent of available social support, and cultural influences.³¹ In addition, it has been observed that the stress-reducing effect of HL is only significant in individuals with low or moderate stress levels, and this effect decreases in individuals with high stress levels.³² These findings indicate that the effect of HL on stress is context-sensitive and does not represent a universal finding.

The results obtained within the scope of the study also reveal that PS levels are negatively related to HLA. As stress levels increase, individuals' awareness of healthy lifestyle behaviors decreases. This finding is consistent with studies indicating that chronic stress negatively affects cognitive functions and leads individuals to prefer short-term comfort-seeking behaviors over long-term health goals.^{33,34} Stress is particularly associated with unhealthy eating habits; individuals under stress tend to make less healthy choices.³⁵ Similarly, it has been found that PS reduces individuals' search for information and awareness about healthy living.¹⁵ In addition, stress has been reported to significantly impact sleep quality, physical activity, and general lifestyle habits.^{14,36,37} On the other hand, some theoretical approaches argue that stress does not always have an adverse effect, but can trigger healthy behaviors when

Table 3. Path analysis results on the mediating role of perceived stress in the effect of health literacy on healthy life awareness

Type	Effect	Estimate	SE	95% CI		β	z	p
				LL	UL			
Indirect	HL $\rightarrow$ PS $\rightarrow$ HLA	0.02	0.01	0.01	0.04	0.02	3.56	<0.001
Component	HL $\rightarrow$ PS	-0.21	0.04	-0.29	-0.14	-0.16	-5.60	<0.001
	PS $\rightarrow$ HLA	-0.11	0.02	-0.15	-0.06	-0.12	-4.61	<0.001
Direct	HL $\rightarrow$ HLA	0.46	0.03	0.40	0.52	0.41	15.23	<0.001
Total	HL $\rightarrow$ HLA	0.48	0.03	0.42	0.54	0.43	16.05	<0.001

SE: Standard error, CI: Confidence interval, LL: Lower limit, UL: Upper limit, HL: Health literacy, PS: Perceived stress, HLA: Healthy living awareness

perceived by the individual as a “challenge”.¹³ In this regard, it should be noted that the impact of stress may vary depending on individual differences, coping strategies, and perception patterns.

One of the study’s most striking findings is that PS partially mediates the relationship between HL and HLA. This mediating effect, tested using structural equation modeling, revealed that as individuals’ HL levels increase, their stress levels decrease, positively impacting HLA. In other words, HL not only influences an individual’s knowledge level but also affects stress levels, thereby supporting HLA in a bidirectional manner. This finding aligns with the theoretical framework of Baron and Kenny’s model, one of the most commonly used frameworks for testing mediation relationships.³⁸ It has been observed that individuals with high HL are better able to evaluate health-related issues and are more resilient in the face of stress; this psychological advantage has been shown to influence healthy lifestyle behaviors positively. This finding is supported by various studies demonstrating that HL is related to knowledge levels and psychological well-being.^{17,39,40} On the other hand, it is emphasized that the effects of HL should be evaluated in terms of environmental factors such as social support, access to health services, and economic conditions.^{41,42} In this context, it should be borne in mind that the indirect effects of HL on stress may be strengthened or weakened by contextual variables.

Overall, this study highlights the need for strategies that increase individuals’ knowledge and support their psychological resilience to improve their awareness of healthy living. In this context, enriching HL education with stress management content could increase individuals’ knowledge and strengthen their psychological resilience. The findings provide strong evidence for developing comprehensive public health policies that support individuals’ health-related decision-making processes. Overall, this study highlights the need for strategies that increase individuals’ knowledge and support their psychological resilience to improve their awareness of healthy living. In this context, enriching HL education with stress management content could increase individuals’ knowledge and strengthen their psychological resilience. The findings provide strong evidence for developing comprehensive public health policies that support individuals’ health-related decision-making processes.

Limitations

This research has certain limitations. Firstly, the data were collected through an online, self-report questionnaire, which may introduce response bias due to social desirability effects. Secondly, the cross-sectional design prevents causal inferences about the relationships among the variables. Thirdly, although the intended target population was the general Turkish public, the overrepresentation of younger and highly educated individuals reduces the representativeness of the findings. Moreover, the use of a convenience sampling method further limits generalizability. Therefore, despite the large sample size, the results should be interpreted cautiously and may not fully reflect the broader Turkish adult population, particularly older adults or those with lower educational attainment.

CONCLUSION

This study demonstrated that health literacy directly increases awareness of healthy living and indirectly reinforces this effect by reducing perceived stress. These findings highlight health literacy not only as an informational resource but also as a psychological asset that supports resilience and well-being. Strengthening health literacy education-particularly when combined with stress management skills-may enhance individuals’ ability to make informed health decisions and adopt healthier lifestyles. For Türkiye, integrating such programs into university curricula and community health initiatives could support national strategies to reduce preventable health risks and promote long-term wellness. Future research should employ longitudinal and experimental designs to confirm causality and explore additional psychosocial variables influencing this relationship.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Selçuk University Faculty of Health Sciences Non-interventional Clinical Researches Ethics Committee (Date: 29.01.2025, Decision No: 2025/50).

Informed Consent

Informed consent was obtained electronically, and participants were allowed access to the questionnaire only after providing their explicit agreement to participate.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- İskender H, Dokumacıoğlu E, Kanbay Y, Kılıç N. Üniversite öğrencilerinde sağlıklı yaşam ve depresyon puan düzeyleri ile ilgili faktörlerin belirlenmesi. *ACU Sağlık Bil Derg.* 2018;9(4):414-423. doi:10.31067/0.2018.64
- Quaye ES, Mokgethi K, Ameyibor LEK. Health self-identity-based motivations and behavioral intentions: a predictive model and segmentation analysis. *Soc Market Quarterly.* 2021;27(4):347-369. doi:10.1177/15245004211053853
- Gayef A. Healthy living behaviors in university students and related factors. *Anatol JFM.* 2019. doi:10.5505/anatoljfm.2018.43534
- Sørensen K, Van den Broucke S, Fullam J, et al. Health literacy and public health: a systematic review and integration of definitions and models. *BMC Public Health.* 2012;12:80. doi:10.1186/1471-2458-12-80
- Yu X, Luo M, Wu S, et al. Developing a questionnaire to evaluate the health information literacy in China. *Front Public Health.* 2023;11:1068648. doi:10.3389/fpubh.2023.1068648

6. Nutbeam D. The evolving concept of health literacy. *Soc Sci Med*. 2008;67(12):2072-2078. doi:10.1016/j.socscimed.2008.09.050
7. Paasche-Orlow MK, Wolf MS. The causal pathways linking health literacy to health outcomes. *Am J Health Behav*. 2007;31(Suppl 1):S19-S26. doi:10.5555/ajhb.2007.31.supp.S19
8. Berkman ND, Sheridan SL, Donahue KE, Halpern DJ, Crotty K. Low health literacy and health outcomes: an updated systematic review. *Ann Intern Med*. 2011;155(2):97-107. doi:10.7326/0003-4819-155-2-201107190-00005
9. Yang P, Ou Y, Yang H, et al. Research on influencing factors and dimensions of health literacy in different age groups: before and after the COVID-19 era in Chongqing, China. *Front Public Health*. 2021;9:690525. doi:10.3389/fpubh.2021.690525
10. Bektas İ, Kudubeş AA, Ayar D, Bektas M. Predicting the healthy lifestyle behaviors of Turkish adolescents based on their health literacy and self-efficacy levels. *J Pediatr Nurs*. 2021;59:e20-e25. doi:10.1016/j.pedn.2021.01.016
11. Elkin N. The relationship between health literacy and healthy lifestyle behaviors: a meta-analysis. *Medicine*. 2024;103(43):e40260. doi:10.1097/MD.00000000000040260
12. Sudore RL, Mehta KM, Simonsick EM, et al. Limited literacy in older people and disparities in health and healthcare access. *J Am Geriatr Soc*. 2006;54(5):770-776. doi:10.1111/j.1532-5415.2006.00691.x
13. Lazarus R, Folkman S. *Stress, Appraisal, and Coping*. Springer US; 1984.
14. Devran H, Beyazit Üçgün A, Yürekli MV, Uskun E. Yaşam kalitesinin yordayıcısı olarak sağlıklı yaşam biçimi davranışları: tıp fakültesi öğretim elemanları örneği. *Türk Halk Sağl Derg*. 2021;19(1):55-68. doi:10.20518/tjph.763355
15. Wang D, Sun X, He F, Liu C, Wu Y. The mediating effect of family health on the relationship between health literacy and mental health: a national cross-sectional survey in China. *Int J Soc Psychiatry*. 2023;69(6):1490-1500. doi:10.1177/00207640231166628
16. Gouin JP, Kiecolt-Glaser JK. The impact of psychological stress on wound healing: methods and mechanisms. *Immun Allergy Clin North Am*. 2011;31(1):81-93. doi:10.1016/j.iac.2010.09.010
17. Hirooka N, Kusano T, Kinoshita S, Nakamoto H. Influence of perceived stress and stress coping adequacy on multiple health-related lifestyle behaviors. *IJERPH*. 2021;19(1):284. doi:10.3390/ijerph19010284
18. Algren MH, Ekholm O, Nielsen L, Ersbøll AK, Bak CK, Andersen PT. Associations between perceived stress, socioeconomic status, and health-risk behaviour in deprived neighbourhoods in Denmark: a cross-sectional study. *BMC Public Health*. 2018;18(1):250. doi:10.1186/s12889-018-5170-x
19. Bandura A. Self-Efficacy: The Exercise of Control. W H Freeman/Times Books/ Henry Holt & Co; 1997:ix, 604.
20. Bozkurt S. Process Makro İle Aracılık, Düzenleyicilik ve Durumsal Aracılık Etki Analizleri (SPSS Uygulamalı). 1. Basım. Ekin Basım Yayın; 2023.
21. Hayes AF. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. Third edition. The Guilford Press; 2022.
22. Pelikan JM, Röthlin F, Ganahl K. Measuring comprehensive health literacy in general populations: Validation of instrument, indices and scales of the HLS-EU study. In: Proceedings of the 6th Annual Health Literacy Research Conference. 2014.
23. Yeşildal M. Validity and reliability of the Turkish version of the HLS-EU-Q6 questionnaire. *BMC Public Health*. 2025;25(1):2633. doi:10.21203/rs.3.rs-5145103/v1
24. Kocapınar F, Ekşi H. Turkish adaptation study of the Perceived Stress Scale short form. In: IKSAD Publications; 2024:554-563.
25. Schäfer SK, Von Boros L, Göritz AS, et al. The Perceived Stress Scale 2&2: a two-factorial German short version of the Perceived Stress Scale. *Front Psychiatry*. 2023;14:1195986. doi:10.3389/fpsy.2023.1195986
26. Özer E, Yılmaz N. Healthy life awareness: a scale development study. *J Tradit Complem Med*. 2020;3(1):47-60. doi:10.5336/jtracom.2019-72138
27. Nutbeam D. Health literacy as a public health goal: a challenge for contemporary health education and communication strategies into the 21st century. *Health Promot Int*. 2000;15(3):259-267. doi:10.1093/heapro/15.3.259
28. Liu Y, Meng H, Tu N, Liu D. The relationship between health literacy, social support, depression, and frailty among community-dwelling older patients with hypertension and diabetes in China. *Front Public Health*. 2020;8:280. doi:10.3389/fpubh.2020.00280
29. Rababah JA, Al-Hammouri MM, Drew BL. The impact of health literacy on college students' psychological disturbances and quality of life: a structural equation modeling analysis. *Health Qual Life Outcomes*. 2020;18(1):292. doi:10.1186/s12955-020-01541-7
30. Pappadis MR, Sander AM, Juengst SB, et al. The relationship of health literacy to health outcomes among individuals with traumatic brain injury: a traumatic brain injury model systems study. *J Head Trauma Rehabil*. 2024;39(2):103-114. doi:10.1097/HTR.0000000000000912
31. Ishikawa H, Yano E. Patient health literacy and participation in the health-care process. *Health Expectations*. 2008;11(2):113-122. doi:10.1111/j.1369-7625.2008.00497.x
32. Zhang H, Li Y, Peng S, Jiang Y, Jin H, Zhang F. The effect of health literacy on COVID-19 vaccine hesitancy: the moderating role of stress. *Vaccine*. 2022;40(32):4473-4478. doi:10.1016/j.vaccine.2022.06.16.21258808
33. Cohen S, Kamarck T, Mermelstein R. A Global measure of perceived stress. *J Health Soc Behav*. 1983;24(4):385. doi:10.2307/2136404
34. McEwen BS. Stress, adaptation, and disease: allostasis and allostatic load. *Ann New York Acad Sci*. 1998;840(1):33-44. doi:10.1111/j.1749-6632.1998.tb09546.x
35. Ng DM, Jeffery RW. Relationships between perceived stress and health behaviors in a sample of working adults. *Health Psychol*. 2003;22(6):638-642. doi:10.1037/0278-6133.22.6.638
36. Ülker İ. Obezitenin hastalık sürecinde yaşam kalitesine etkisi: COVID-19 örneği. *Adnan Menderes Üniv Sağ Bil Fak Derg*. 2023;7(3):682-691. doi:10.46237/amusbd.1338547
37. Yeşil Bayülgen M, Altıok M. Healthy lifestyle behaviours of percutaneous transluminal coronary angioplasty patients and the affecting factors. *J Cardiovasc Nurs*. 2017;8(16):45-54. doi:10.5543/khd.2017.28199
38. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Person Soc Psychol*. 1986;51(6):1173-1182. doi:10.1037/0022-3514.51.6.1173
39. Angermayr L, Melchart D, Linde K. Multifactorial lifestyle interventions in the primary and secondary prevention of cardiovascular disease and type 2 diabetes mellitus-a systematic review of randomized controlled trials. *Ann Behav Med*. 2010;40(1):49-64. doi:10.1007/s12160-010-9206-4
40. McKenzie SH, Harris MF. Understanding the relationship between stress, distress and healthy lifestyle behaviour: a qualitative study of patients and general practitioners. *BMC Fam Pract*. 2013;14(1):166. doi:10.1186/1471-2296-14-166
41. Gazmararian JA. Health literacy among medicare enrollees in a managed care organization. *JAMA*. 1999;281(6):545. doi:10.1001/jama.281.6.545
42. Jordan TR, Khubchandani J, Wimblishauser M. The impact of perceived stress and coping adequacy on the health of nurses: a pilot investigation. *Nurs Res Pract*. 2016;2016:1-11. doi:10.1155/2016/5843256

Evaluation of carotid artery calcification on panoramic radiographies

 Ali Altındağ¹,  Ömer Altındağ²

¹Department of Dentomaxillofacial Radiology, Faculty of Dentistry, Necmettin Erbakan University, Konya, Türkiye

²Department of Arts and Sciences, Statistics and Computer Science, Faculty of Arts and Sciences,
Bilecik Şeyh Edebali University, Bilecik, Türkiye

Cite this article as: Altındağ A, Altındağ Ö. Evaluation of carotid artery calcification on panoramic radiographies. *J Health Sci Med.* 2025;8(6):1006-1010.

Received: 20.08.2025

Accepted: 18.09.2025

Published: 25.10.2025

ABSTRACT

Aims: The aim of this study was to evaluate the prevalence of carotid artery calcification (CAC) on panoramic radiographs and to investigate its relationship with age, sex, and systemic diseases.

Methods: Panoramic radiographs obtained from 599 patients who referred to our clinic were retrospectively analyzed. Irregular, heterogeneous radiopaque masses located at the C3-C4 cervical vertebral level were considered indicative of CAC, with differential diagnosis performed against other calcified anatomical or pathological structures. Data on patient age, sex, and systemic conditions (cardiovascular disease, diabetes) were recorded. Information on smoking status, lipid profiles, medication use, and history of cardiovascular surgery was not recorded. Chi-square tests and logistic regression analyses were used to assess associations between CAC and clinical variables.

Results: CAC was detected in 19 patients, corresponding to a prevalence of 3.2%. The prevalence increased with age: 0.9% in the 18-35 age group, 6.3% in the 36-60 age group, and 4.9% in patients aged ≥ 61 years. Logistic regression analysis identified age as an independent risk factor for CAC ($p < 0.05$). No significant differences were observed with respect to sex or systemic diseases; however, the prevalence of CAC was relatively higher in patients with cardiovascular disease (7.7%), whereas no cases were observed among diabetics.

Conclusion: The prevalence of CAC detected incidentally on panoramic radiographs was within the range reported in previous studies. Age was confirmed as an independent risk factor, while sex and systemic comorbidities showed no significant associations. Panoramic radiographs may serve as a useful tool for the incidental detection of CAC in routine dental practice.

Keywords: Atherosclerosis, carotid arteries, diagnostic imaging, panoramic radiography

INTRODUCTION

Calcification is defined as the abnormal deposition of calcium salts in tissues and organs.¹ When this deposition results in an organized structure within soft tissues, it is termed heterotopic ossification; when the deposition occurs without organization, it is referred to as heterotopic calcification. Soft tissue calcifications are rare in children but are more frequently observed in individuals over the age of 40. In the head and neck region, calcifications are commonly seen in the laryngeal cartilages, vertebrae, arteries, and thyroid gland.^{2,3}

Calcification and plaque formation in the carotid arteries may lead to carotid artery calcification (CAC). Atherosclerosis-related calcifications are most often observed in the internal carotid artery.^{4,5} Hypertension, elevated serum cholesterol and carbohydrate levels, and systemic disorders that may impair endothelial integrity are among the main etiological factors of CAC.⁶

Atherosclerosis is a progressive inflammatory disease and represents a major cause of morbidity and mortality in

developed countries. It has been reported to lead to peripheral arterial disease, coronary artery disease, and stroke.^{7,8} The most commonly affected site is the internal carotid artery, and approximately 40% of cerebrovascular events caused by atherosclerotic plaques are reported in this region.⁹ According to global data, cerebrovascular events such as stroke constitute a leading cause of death and disability worldwide; specifically, stroke ranks as the second leading cause of death and the third leading cause of death and disability combined (DALYs) globally and across regions including North America, Europe, and Asia.^{10,11}

In dentistry, panoramic radiographs are widely used along with intraoral radiographs for diagnosis, treatment planning, and post-treatment evaluation.^{3,12,13} The main advantage of panoramic radiographs is their ability to visualize the teeth, alveolar bones, and anatomical structures of the head and neck with a lower radiation dose compared to other imaging modalities.^{14,15} Since early detection of risk factors is crucial

Corresponding Author: Ali Altındağ, aaltindag@erbakan.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

for reducing morbidity and mortality, imaging plays an important role in preventive strategies.¹⁶

The use of panoramic radiographs for the detection of CAC was first highlighted by Friedlander and Lande in 1981.¹⁷ They reported that CAC, which can be observed at the C3-C4 vertebral level and is a major cause of stroke, may be identified on panoramic radiographs.¹⁷ On panoramic films, these lesions typically appear at the C3-C4 level as irregularly bordered, heterogeneous radiopaque formations (Figure).^{3,17,18} Thus, panoramic radiography not only provides a rapid and cost-effective evaluation of dental and maxillofacial structures but also enables the incidental detection of CAC.¹⁹



Figure. Bilateral carotid artery calcifications at the C3-C4 vertebral level on panoramic radiograph (indicated by red arrows)

The present study aimed to determine the prevalence of carotid artery calcifications incidentally detected on panoramic radiographs, which may occasionally be overlooked, and to evaluate their distribution according to age, sex, and systemic disease status.

METHODS

Ethical Approval

This study was approved by the Necmettin Erbakan University Dentistry Non-drug and Non-medical Device Researches Ethics Committee (Date: 31.03.2022, Decision No: 2022/16-120), and conducted in accordance with the principles of the Declaration of Helsinki.

Study Design

Panoramic radiographs of patients who presented to the Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Necmettin Erbakan University, for various reasons between January 2020 and October 2021 were retrospectively reviewed and randomly selected. All panoramic radiographs were obtained using a Morita Veraviewepocs 2D panoramic unit (J. Morita MFG Corp., Kyoto, Japan) at 60-70 kVp, 5-7 mA, and 6-8 s exposure time, according to the manufacturer's recommendations.

Inclusion criteria:

- Individuals aged 18 years and older
- Patients with complete demographic and clinical records
- Panoramic radiographs of diagnostically acceptable image quality

Exclusion criteria:

- Patients with a history of head and neck surgery
- Radiographs with artifacts, positioning errors, or pathologies reducing diagnostic quality
- Patients with incomplete demographic data

Data on patient age, sex, and systemic conditions were collected from clinical records. Systemic diseases included cardiovascular disorders (e.g., hypertension, ischemic heart disease, arrhythmia, congenital heart disease) and diabetes mellitus. Data regarding patients' smoking habits, lipid profiles (LDL cholesterol and triglycerides), and medication use (e.g., antihypertensives, anticoagulants, antiplatelet agents, lipid-lowering drugs) were not recorded. Information about previous cardiovascular interventions (e.g., bypass surgery, angioplasty, stenting) was also not recorded.

Of the 750 radiographs screened, 599 were included in the study.

A priori power analysis was performed to determine the minimum required sample size for detecting differences in the prevalence of carotid artery calcification (CAC) between age groups. Using a two-tailed comparison of two independent proportions with $\alpha=0.05$ and power=0.80, the required total sample size was estimated to be 312. In addition, when the age groups were analyzed together in a 3×2 contingency table (18-35, 36-60, ≥61 years vs. presence/absence of CAC), the effect size (Cohen's w) was 0.15, and the minimum required total sample size was approximately 443 to achieve 80% power at $\alpha=0.05$. As our study included 599 individuals, the achieved power was sufficient for the planned analyses.

At the cervical C3-C4 level, irregular, heterogeneous, vertically linear, or nodular radiopaque masses were recorded as CAC. In the differential diagnosis, anatomical structures such as thyroid cartilage, hyoid bone, calcified stylohyoid ligament, and epiglottis, as well as calcified lymph nodes, phleboliths, submandibular salivary gland stones, and tonsilloliths were considered.²⁰

Statistical Analysis

Data were analyzed using SPSS software (Version 21.0; IBM Corp., Armonk, NY, USA). Descriptive statistics (mean±SD, median, frequency, and percentage) were calculated. Associations between categorical variables were assessed using the chi-square test. For non-normally distributed age data, the Mann-Whitney U test was applied.

To identify independent risk factors for CAC, logistic regression analysis was performed. All results were evaluated at a 95% confidence interval with a significance level set at $p<0.05$.

To assess intraobserver reliability, 100 randomly selected radiographs were re-evaluated by the same examiner after 20 days, and Cohen's kappa coefficient was calculated.

RESULTS

Intraobserver agreement was excellent, with a Cohen's kappa coefficient of 0.97, indicating high reliability in the detection of CAC.

A total of 599 individuals were included in this study. The mean age of the participants was 36.1±13.8 years (range: 18-84), with a median age of 33 years. Of the participants, 74.5% were female (n=446) and 25.5% were male (n=153). A total of 18.4% (n=110) reported a systemic disease, including cardiovascular disease (n=26, 4.3%) and diabetes mellitus (n=20, 3.3%). The cardiovascular disease group comprised heterogeneous conditions, predominantly hypertension, ischemic heart disease, and arrhythmias. Carotid artery calcification (CAC) was detected in 3.2% (n=19) of all participants.

When stratified by age groups, CAC was detected in 3 of 335 patients (0.9%) in the 18-35 age group, 14 of 223 patients (6.3%) in the 36-60 age group, and 2 of 41 patients (4.9%) in those aged 61 years and older. The difference between age groups was statistically significant (p=0.001). No significant association was observed between CAC and sex (female: 3.4% vs. male: 2.6%, p=0.379). Similarly, the presence of systemic disease (p=0.226), cardiovascular disease (p=0.440), or diabetes mellitus (p=0.862) did not show a statistically significant relationship with CAC (Table 1). However, given the limited number of participants with cardiovascular disease (n=26), these findings should be interpreted with caution.

Table 1. Distribution of carotid artery calcification according to age groups, sex, and systemic disease subgroups				
Parameter	n (%)	CAC present n (%)	CAC absent n (%)	p-value
Sex				
Female	446 (74.5)	15 (78.9)	431 (74.3)	0.379
Male	153 (25.5)	4 (21.1)	149 (25.7)	
Age groups				
18-35	335 (55.9)	3 (15.8)	332 (57.2)	0.001
36-60	223 (37.2)	14 (73.7)	209 (36.0)	
≥61	41 (6.8)	2 (10.5)	39 (6.7)	
Systemic disease				
Absent	489 (81.6)	13 (68.4)	476 (82.1)	0.226
Present	110 (18.4)	6 (31.6)	104 (17.9)	
Cardiovascular disease				
Absent	573 (95.7)	17 (89.5)	556 (95.9)	0.440
Present	26 (4.3)	2 (10.5)	24 (4.1)	
Diabetes mellitus				
Absent	579 (96.7)	19 (100.0)	560 (96.6)	0.862
Present	20 (3.3)	0 (0.0)	20 (3.4)	
Values are presented as numbers and percentages. Statistical comparisons were performed using the Chi-square test. A p-value of <0.05 was considered statistically significant. CAC: Carotid artery calcification				

In logistic regression analysis, age was identified as the only independent risk factor for CAC. Each one-year increase in age was associated with approximately a 4% increase in the likelihood of developing CAC (OR=1.04; 95% CI: 1.01-1.08; p=0.014). Sex (OR=1.53; p=0.402), systemic disease (OR=1.30; p=0.684), cardiovascular disease (OR=1.12; p=0.900), and diabetes (OR=0.00; p=0.993) were not significant independent predictors (Table 2).

Table 2. Logistic regression analysis of risk factors associated with carotid artery calcification			
Variable	OR	95% CI (lower-upper)	p-value
Age (years)	1.04	1.01-1.08	0.014**
Sex (male)	1.53	0.56-4.18	0.402
Systemic disease	1.30	0.36-4.67	0.684
Cardiovascular disease	1.12	0.18-7.11	0.900
Diabetes mellitus	0.00	0.00-∞	0.993
Values are presented as OR with 95% CI. Logistic regression analysis was applied to evaluate independent associations between variables and CAC. A p-value of <0.05 was considered statistically significant. Odds ratio, CI: Confidence interval, CAC: Carotid artery calcification			

Subgroup analyses revealed that, among participants with systemic disease, the prevalence of CAC was highest in the 36-60 age group (6.7%), but the differences across age groups were not statistically significant (p=0.61). When sex and age were evaluated together, a significant difference in CAC prevalence was observed (p=0.012), with the highest rate detected in men aged ≥61 years (9.1%) compared to women in the same age group (3.3%). In participants with cardiovascular disease, the prevalence of CAC was higher (7.7%) than in those without (3.0%), although the difference was not statistically significant (p=0.440). No cases of CAC were observed among participants with diabetes mellitus (0/20).

Overall, the prevalence of CAC in this population was 3.2%, with age emerging as the strongest independent risk factor. Other variables, including sex, systemic disease, cardiovascular disease, and diabetes, were not significantly associated with CAC.

DISCUSSION

Atherosclerotic cardiovascular diseases encompass a group of conditions that include fatal and nonfatal coronary artery diseases (such as myocardial infarction and angina), ischemic cerebrovascular diseases (stroke), and peripheral arterial diseases. Panoramic radiographs, which are routinely obtained during dental treatment, may be useful in detecting CAC, a known risk factor for stroke.²¹ Dentists are able to distinguish CAC from anatomical and pathological structures in the differential diagnosis.²² Considering the health benefits of early diagnosis and treatment, the prevalence of CAC was evaluated on panoramic radiographs of patients who presented to our clinic for various dental reasons. The early detection of CAC by dentists is regarded as an important step toward reducing the risk of stroke.

In this study, the prevalence of carotid artery calcification (CAC) incidentally detected on panoramic radiographs was 3.2%. Previous studies have reported that the prevalence of CAC varies across populations depending on sample characteristics and diagnostic criteria, but it is generally within the range of 2% to 5%. Higher rates have been observed particularly in older individuals and in those with systemic comorbidities.^{23,24} Our findings therefore fall within this expected range, supporting the notion that panoramic radiography can reveal clinically relevant incidental findings and are in agreement with prior studies conducted in dental and general populations.

In a study by Hubar et al.,²⁵ which evaluated 700 patients, the prevalence of CAC was reported as 0.43%. However, the individuals included in that study were between 14 and 77 years of age, whereas in the present study the age range was 18 to 84 years. The discrepancy between the two studies may be attributed to differences in age distribution and racial background.²⁵ In a prevalence study conducted among Japanese individuals over the age of 80, the rate of CAC was found to be 5%.²⁶ Furthermore, when comparing smokers by sex, men were reported to have a higher risk of stroke and to be at greater risk compared to non-smokers.²⁷ The lack of assessment of lifestyle factors such as smoking and diet in our study should therefore be considered among its limitations.

According to the findings of our study, logistic regression analysis identified age as an independent risk factor for CAC. The increase in CAC prevalence with advancing age has been consistently emphasized in large-sample studies and systematic reviews.^{24,28} In particular, individuals over the age of 50 have been reported to demonstrate a markedly higher risk.

In terms of sex, no significant difference was observed in our study. The literature, however, presents conflicting results: some investigations have reported a higher prevalence of CAC in men, while others have highlighted an increased risk among postmenopausal women. For instance, Almog et al.²⁰ found higher rates in men, although the difference did not reach statistical significance. Conversely, other series have suggested that male sex may act as an independent risk factor.²⁹ These discrepancies imply that the influence of sex on CAC may vary according to population characteristics and age distribution. Our findings therefore appear to be more consistent with studies that did not identify sex as an independent determinant of CAC.

With respect to systemic conditions, no statistically significant association was found between the presence of CAC and either cardiovascular disease or diabetes. Nevertheless, the prevalence of CAC was relatively higher among patients with cardiovascular disease (7.7%), whereas no cases were identified among diabetic individuals. The literature also presents divergent findings on this issue: while some studies have reported significant associations between CAC and cerebrovascular events, coronary artery disease, diabetes, and hypertension,^{30,31} others have failed to confirm such relationships.³¹ The absence of significant associations in our study may, at least in part, be attributed to the small sample size within these subgroups. Future studies with larger cohorts and a more balanced distribution of systemic comorbidities are warranted to clarify these relationships.

From a diagnostic perspective, panoramic radiographs, in addition to their widespread use in dentistry, provide the opportunity to incidentally detect clinically significant findings such as CAC. Friedlander and Lande¹⁷ were the first to describe the presence of CAC on panoramic radiographs and suggested that this method could be useful in identifying individuals at increased risk of stroke.¹⁰ Subsequent studies have confirmed that panoramic radiographs possess a certain diagnostic value in detecting CAC, although confirmation by complementary methods-particularly Doppler

ultrasonography and, when indicated, cone-beam computed tomography (CBCT)-has been recommended.³²⁻³⁴ Therefore, when radiographic findings consistent with CAC are observed on panoramic images, referral of patients to the appropriate medical units is of critical importance. This further underscores the potential role of dentists in interdisciplinary prevention strategies aimed at reducing stroke risk.

Limitations

The limitations of this study include its single-center and cross-sectional design, the use of a single observer for image evaluation, and the small number of cases within certain systemic disease subgroups. Additionally, data regarding patients' smoking habits, lipid profiles (LDL, triglycerides), and medication use (e.g., antihypertensives, anticoagulants, and antiplatelet agents) were not recorded. Furthermore, the cardiovascular disease group was heterogeneous, consisting of different conditions such as hypertension, ischemic heart disease, and arrhythmias, which may have influenced the results. These factors, which may influence the presence of CAC, should be addressed in future studies. Another important limitation is the absence of confirmation with Doppler ultrasonography, which is considered the gold standard for CAC assessment. Larger, multicenter, and prospective studies are therefore needed to more reliably elucidate the relationship between CAC detected on panoramic radiographs and clinical outcomes.

CONCLUSION

In this retrospective study, CAC was identified in 3.2% of panoramic radiographs, with prevalence increasing significantly with age. Logistic regression confirmed age as an independent risk factor for CAC, consistent with findings from previous epidemiological studies. Although no statistically significant associations were found with sex or systemic comorbidities, a higher prevalence was observed in patients with cardiovascular disease. Panoramic radiographs, while primarily used for dental purposes, may thus provide valuable information regarding systemic vascular health. The incidental detection of CAC highlights the important role of dentists in interdisciplinary prevention strategies and supports the referral of suspected cases for further medical evaluation to reduce stroke risk.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Necmettin Erbakan University Dentistry Non-drug and Non-medical Device Researches Ethics Committee (Date: 31.03.2022, Decision No: 2022/16-120).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Mousa AI, Hesham NM, Hamza SA, Mohamed EM. The prevalence of carotid artery calcification detected from the panoramic radiographs of a sample of the Egyptian population: a retrospective study. *Egypt Dent J*. 2025;71:2135-2143. doi:10.21608/edj.2025.373220.3437
- Özemre MÖ, Seçgin CK, Gülşahı A. Yumuşak doku kalsifikasyonları ve ossifikasyonları: derleme. *Acta Odontol Turcica*. 2016;33:166-175. doi:10.17214/aot.55965
- Altındağ A. Investigation of calcified atherosclerotic plaques on panoramic radiographs. *Ann Med Res*. 2024;31(1):66-70. doi:10.5455/annalsmedres.2023.12.323
- Aydinyurt HŞ, Taşkın C. Doğu Anadolu Bölgesi'nde karotis arter kalsifikasyonlarının sıklığının belirlenmesi. *Van Dent J*. 2020;1(1):1-6.
- Altındağ A, Yüksel İ. Panoramik radyografide yumuşak doku kalsifikasyon ve ossifikasyonları. *Necmettin Erbakan Uni Dent J*. 2021;3(2):62-69. doi:10.51122/neudentj.2021.18
- Yıldırım D, Bilgir E. Baş boyun bölgesindeki yumuşak doku kalsifikasyon ve ossifikasyonları. *Atatürk Univ Dis Hek Fak Derg*. 2015;25(13):82-90. doi:10.17567/dfd.53076
- Ross R. Atherosclerosis-an inflammatory disease. *N Engl J Med*. 1999;340:115-126. doi:10.1056/NEJM199901143400207
- Gusev E, Sarapultsev A. Atherosclerosis and inflammation: insights from the theory of general pathological processes. *Int J Mol Sci*. 2023;24(9):7910. doi:10.3390/ijms24097910
- Yang SS, Woo SY, Kim DI. Analysis of atherosclerotic plaque distribution in the carotid artery. *Clin Cardiol*. 2022;45(12):1272-1276. doi:10.1002/clc.23903
- Global Stroke Collaboration, Neurology. 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(10):795. doi:10.1016/S1474-4422(21)00252-0
- Feigin VL, Brainin M, Norrving B, et al. World stroke organization: global stroke fact sheet 2025. *Int J Stroke*. 2025;20(2):132-144. doi:10.1177/17474930241308142
- Tugnait A, Clerehugh D, Hirschmann P. Survey of radiographic practices for periodontal disease in UK and Irish dental teaching hospitals. *Dentomaxillofac Radiol*. 2000;29(6):376-381. doi: 10.1038/sj/dmfr/4600561
- Balkan EP, Deniz HA, Kurt MH, Samunahmetoğlu E, Karahan S. Application of fractal analysis in detecting trabecular bone characteristics around mandibular impacted third molars on dental panoramic radiographs. *Eur Ann Dent Sci*. 2024;51(2):67-73. doi:10.52037/eads.2024.0008
- Izzetti R. Basic knowledge and new advances in panoramic radiography imaging techniques: a narrative review on what dentists and radiologists should know. *App Sci*. 2021;11(17):7858.
- Altındağ A, Cora AH. Evaluation of maxillofacial soft tissue calcifications in edentulous patients on panoramic radiographs. *Geriatrisk Bil Derg*. 2023;6(2):103-113. doi:10.47141/geriatrisk.1292936
- Schroder AGD, de Araujo CM, Guariza-Filho O, Flores-Mir C, de Luca Canto G, Porporatti AL. Diagnostic accuracy of panoramic radiography in the detection of calcified carotid artery atheroma: a meta-analysis. *Clin Oral Invest*. 2019;23(5):2021-2040. doi:10.1007/s00784-019-02880-6
- Friedlander AH, Lande AJ. Panoramic radiographic identification of carotid arterial plaques. *Oral Surg Oral Med Oral Pathol*. 1981;52(1):102-104. doi:10.1016/0030-4220(81)90181-X
- Sari AH, Mağat G. Prevalence of carotid artery calcification in hemodialysis patients: a case-control study. *Van Dent J*. 2025;6(1):1-13.
- Ghassemzadeh S, Sbricoli L, Frigo AC, Bacci C. Incidental findings detected with panoramic radiography: prevalence calculated on a sample of 2017 cases treated at a major Italian trauma and cancer centre. *Oral Radiol*. 2021;37(3):507-517. doi:10.1007/s11282-020-00488-1
- Almog DM, Tsimidis K, Moss ME, Gottlieb RH, Carter LC. Evaluation of a training program for detection of carotid artery calcifications on panoramic radiographs. *Oral Surg Oral Med Oral Pathol*. 2000;90(1):111-117. doi:10.1067/moe.2000.107056
- Friedlander A. Identification of stroke-prone patients by panoramic and cervical spine radiography. *Dentomaxillofac Radiol*. 1995;24(3):160-164. doi:10.1259/dmfr.24.3.8617388
- Kumagai M, Yamagishi T, Fukui N, Chiba M. Carotid artery calcification seen on panoramic dental radiographs in the Asian population in Japan. *Dentomaxillofac Radiol*. 2007;36(2):92-96. doi:10.1259/dmfr/79378783
- Soares GC, Kurita LM. Prevalence of carotid artery calcifications among 2,500 digital panoramic radiographs of an adult Brazilian population. *Med Oral Patol Oral Cir Bucal*. 2018;23(3):e256. doi:10.4317/medoral.22350
- Janiszewska-Olszowska J, Jakubowska A, Gieruszczak E, Jakubowski K, Wawrzyniak P, Grocholewicz K. Carotid artery calcifications on panoramic radiographs. *Int J Environ Res Public Health*. 2022;19(21):14056. doi:10.3390/ijerph192114056
- Hubar J. Carotid artery calcification in the black population: a retrospective study on panoramic radiographs. *Dentomaxillofac Radiol*. 1999;28(6):348-350. doi:10.1038/sj/dmfr/4600478
- Ohba T, Takata Y, Ansai T, et al. Evaluation of calcified carotid artery atheromas detected by panoramic radiograph among 80-year-olds. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2003;96(5):647-650. doi:10.1016/j.tripleo.2003.07.001
- Peters SA, Huxley RR, Woodward M. Smoking as a risk factor for stroke in women compared with men: a systematic review and meta-analysis of 81 cohorts, including 3,980,359 individuals and 42,401 strokes. *Stroke*. 2013;44(10):2821-2828. doi:10.1161/STROKEAHA.113.002342
- Mai V, Taneja A, Larjava H, Chehroudi B, MacDonald D, Luong C. Calcified carotid artery atheroma on standard dental radiographs: a public health opportunity for cardiovascular risk reduction. *Am J Prevent Cardiol*. 2024;19:100714. doi:10.1016/j.ajpc.2024.100714
- Brar A, DeColibus K, Rasner DS, et al. Carotid artery calcification detected on panoramic radiography is significantly related to cerebrovascular accident, coronary artery disease, and poor oral health: a retrospective cross-sectional study. *Dentistry J*. 2024;12(4):99. doi:10.3390/dj12040099
- Gustafsson N, Ahlqvist J, Norhammar A, et al. Association of high cardiovascular risk and diabetes with calcified carotid artery atheromas depicted on panoramic radiographs. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2022;133(1):88-99. doi:10.1016/j.oooo.2021.06.006
- Alves N, Deana NF, Garay I. Detection of common carotid artery calcifications on panoramic radiographs: prevalence and reliability. *Int J Clin Exp Med*. 2014;7(8):1931-1939.
- Khosropanah S, Shahidi S, Bronoosh P, Rasekhi A. Evaluation of carotid calcification detected using panoramic radiography and carotid Doppler sonography in patients with and without coronary artery disease. *Br Dent J*. 2009;207(4):E8. doi:10.1038/sj.bdj.2009.762
- Bladh M, Gustafsson N, Engström G, et al. Defined shapes of carotid artery calcifications on panoramic radiographs correlate with specific signs of cardiovascular disease on ultrasound examination. *Oral Surg Oral Med Oral Pathol Oral Radiol*. 2024;137(4):408-420. doi:10.1016/j.oooo.2023.12.783
- Carter LC, Tsimidis K, Fabiano J. Carotid calcifications on panoramic radiography identify an asymptomatic male patient at risk for stroke: a case report. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 1998;85(1):119-122. doi:10.1016/S1079-2104(98)90409-7

Evaluation of sex estimation through mandibular morphometric measurements on cone-beam computed tomography images

Elif Çelebi¹, Sema Nur Ersoy Kölege², Özgür Koşkan³, Özlem Görmez²

¹Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Bahçeşehir University, İstanbul, Türkiye

²Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Süleyman Demirel University, Isparta, Türkiye

³Department of Biometry, Faculty of Agriculture, Süleyman Demirel University, Isparta, Türkiye

Cite this article as: Çelebi E, Ersoy Kölege SN, Koşkan Ö, Görmez Ö. Evaluation of sex estimation through mandibular morphometric measurements on cone-beam computed tomography images. *J Health Sci Med.* 2025;8(6):1011-1017.

Received: 28.07.2025

Accepted: 21.09.2025

Published: 25.10.2025

ABSTRACT

Aims: Sex estimation of unidentified human remains is a critical component of forensic investigations. Among facial bones, the mandible is the most robust and resilient structure, making it a valuable element for personal identification. The aim of this study was to evaluate the potential of mandibular metric parameters-obtained via cone-beam computed tomography (CBCT)-for determining sex.

Methods: CBCT scans of 199 individuals (72 males and 127 females), aged between 18 and 81 years, were retrospectively analyzed. Nine mandibular parameters were measured including the mandibular angle, coronoid and condylar lengths, minimal ramus breadth, and mandibular length, along with bigonial breadth, bicondylar breadth, bimental breadth, and symphyseal height. Statistical analyses included repeated measures ANOVA, independent samples t-test, Pearson correlation, and discriminant function analysis.

Results: Males consistently exhibited higher mean values across all parameters, except for the mandibular angle, which was slightly greater in females (female: 119.33 ± 0.56 ; male: 118.79 ± 0.75) but not statistically significant ($p=0.566$). The most influential variables, based on structure matrix coefficients, were left coronoid length (0.791), right coronoid length (0.649), left condylar length (0.492), and bigonial breadth (0.477). Discriminant function analysis yielded a high overall sex classification accuracy of 91.5%.

Conclusion: The mandible demonstrated high discriminative power, with CBCT-derived measurements providing accurate and reproducible indicators of sexual dimorphism. These findings underscore the utility of mandibular morphometry as a reliable tool for sex estimation in forensic practice.

Keywords: Forensic sciences, cone beam computed tomography, sex determination by skeleton, anthropometry

INTRODUCTION

In recent years, there has been an increase in the number of unidentified human remains worldwide due to various disasters, including fires, plane crashes, natural calamities, building collapses, traffic accidents, homicides, and terrorist attacks. Therefore, establishing the identity of a deceased individual is crucial from legal, criminal, and ethical perspectives.¹ Personal identification requires interdisciplinary knowledge and expertise, particularly in fields such as anatomy, radiology, archaeology, and dentistry. The first and most essential step in the identification process is sex determination, as it immediately eliminates approximately half of the population and provides the basis for further estimations of age and stature using sex-specific standards.^{2,3}

Sexual dimorphism plays a critical role in narrowing down identity by guiding the identification process toward a more accurate conclusion. Traditionally, bones such as the pelvis, skull, and long bones have been used for sex determination.

Among these, the skull is considered the most sexually dimorphic structure after the pelvis. Since craniofacial structures are predominantly composed of hard tissues, they are highly valuable in sex estimation.¹ The mandible, as one of the most robust, largest, and morphologically variable bones of the skull, is particularly useful. Its dense cortical bone structure provides exceptional preservation against trauma, making it an ideal candidate for forensic analysis. Sexual dimorphism in the mandible is primarily reflected in its size and shape, with male mandibles generally being larger and more robust than those of females.^{1,4-6}

Previous osteometric studies have employed both dry bones and radiographic images for sex estimation. Cone-beam computed tomography (CBCT) has been shown to provide higher accuracy, reproducibility, and reliability compared to conventional methods. CBCT produces high-resolution, undistorted three-dimensional images with relatively low

Corresponding Author: Elif Çelebi, celebielifin@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

radiation exposure and is non-invasive, preserving the integrity of the examined tissues.^{5,7,8} Several studies have assessed sexually dimorphic mandibular morphometric parameters such as mandibular angle, ramus height, minimal ramus breadth, bigonial breadth, and bicondylar breadth using CBCT images.^{4,5,9} In addition to using measurement values, landmarks plotting on the 3D model surfaces of the mandible has been reported in the literature for sex estimation.^{10,11} At the same time, a growing body of research has explored Artificial Intelligence (AI) methods, including machine learning algorithms,^{12,13} convolutional neural networks (CNNs),¹⁴ and multi-task learning frameworks,¹⁵ which further emphasize the future potential of automated forensic identification systems.

The aim of the present study was to evaluate mandibular metric parameters obtained from CBCT images in a Turkish subpopulation and to determine their predictive accuracy for sex estimation. Unlike many previous studies, it uses a relatively large sample size and multiple bilateral parameters to provide robust discriminant models applicable to forensic identification. The hypothesis of the study is that CBCT-derived mandibular measurements would exhibit significant sexual dimorphism and that a discriminant function based on these parameters would yield high accuracy in sex classification.

METHODS

This retrospective study was conducted using CBCT images obtained between 2016 and 2024 at the Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Süleyman Demirel University. Ethical approval was granted by the Süleyman Demirel University Health Sciences Ethics Committee (Date: 04.07.2025, Decision No: 35). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki. A total of 238 CBCT scans taken for various clinical indications were initially screened. All CBCT examinations were performed on the same device (Planmeca ProMax 3D Mid, Helsinki, Finland) under standardized acquisition parameters, with a 20×10 cm field of view and a voxel size of 0.4 mm. After applying inclusion and exclusion criteria, 199 healthy individuals were included in the study. Individuals aged 18 years or older with complete visualization of the mandible within the field of view were considered eligible.

Subjects with a history of congenital anomalies, trauma, tumors, or surgical interventions involving the jawbones were excluded. A priori power analysis was performed using G*power software (version 3.1.9.4, University of Düsseldorf, Germany). Assuming a medium effect size (Cohen's $d=0.5$), with a significance level of 0.05 and a statistical power of 80%, the minimum required sample size was calculated as 66 participants. Demographic data was retrieved from the institutional Picture Archiving and Communication System (PACS). CBCT images were obtained from Romexis software version 3.8.3 (Planmeca, Helsinki, Finland). Mandibular linear and angle measurements were performed using QuickVision 3D software. Anatomical landmarks were localized on three dimensional reconstructions and confirmed in axial, coronal and sagittal planes. A single trained investigator performed all measurements to avoid inter-observer variability. For reliability assessment, 40 CBCT scans were re-evaluated, and intraclass correlation coefficients (ICC) were calculated. Definitions, references, and measurement planes for each mandibular parameter are summarized in Table 1, with sample CBCT representations shown in Figure 1 and 2.

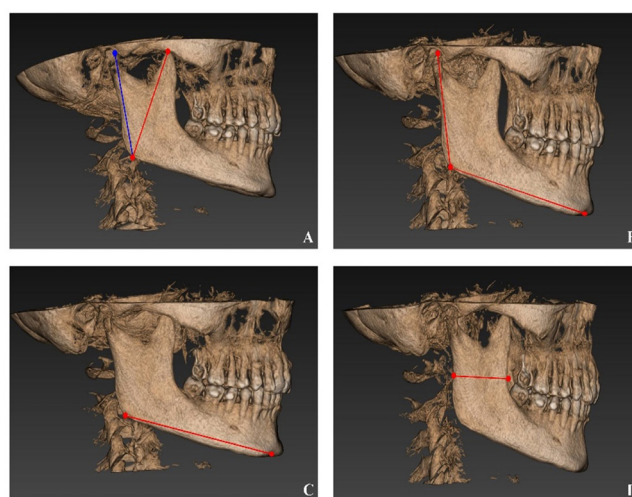


Figure 1. Representative CBCT three-dimensional reconstructions (lateral view) demonstrating the linear and angular measurements performed in the study. (A) Coronoid length (blue line), defined as the distance between the coronoid process tip and gonion; and condylar length (red line), defined as the distance between condylion and gonion. (B) Mandibular angle, measured as the angle formed between the condylion-gonion and menton-gonion lines. (C) Mandibular length, measured as the linear distance from menton to gonion. (D) Minimum ramus breadth, defined as the narrowest width of the mandibular ramus in the sagittal view

CBCT: Cone-beam computed tomography

Table 1. Definitions, references, and cone beam computed tomography reconstruction planes used for the measurement of mandibular parameters

Parameter	Definition	Measurement plane	Reference
Coronoid length	Distance between the coronoid and gonion points	Sagittal view	Assari et al., ⁵
Condylar length	Distance between the condylion and gonion points	Sagittal view	Assari et al., ⁵
Mandibular length	Distance between the menton and gonion points	Sagittal view	Assari et al., ⁵
Mandibular angle	Angle between condylion-gonion and menton-gonion lines	Sagittal view	Assari et al., ⁵
Minimum ramus breadth	Narrowest width of the mandibular ramus	Sagittal view	Kallalli et al., ⁴
Bigonial breadth	Distance between the right and left gonion points	Axial view	Kallalli et al., ⁴
Bicondylar breadth	Distance between the most lateral points of the condyles	Axial view	Kallalli et al., ⁴
Bimental breadth	Distance between the mental foramina	Coronal view	Kumar and Lokanadham, ¹⁶
Symphyseal height	Distance from the alveolar crest of the mandibular central incisors to the inferior mandibular border	Coronal view	Direk et al. ¹⁷

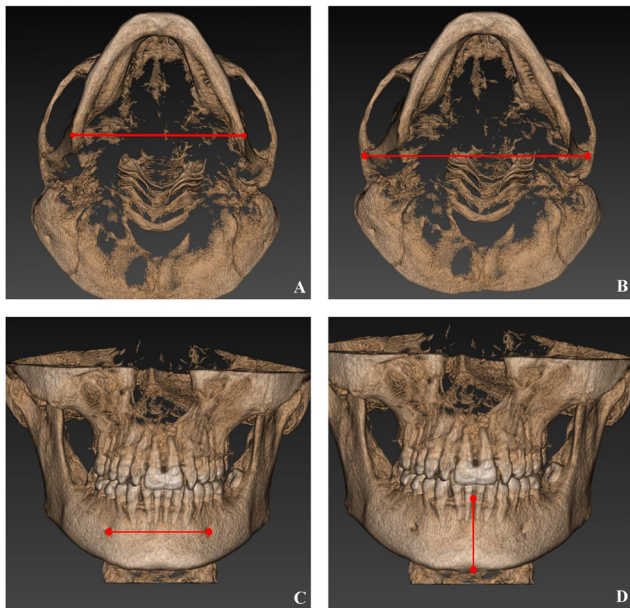


Figure 2. Representative CBCT three-dimensional reconstructions (axial and coronal views) illustrating transverse and vertical mandibular measurements. (A) Bigonial breadth, defined as the distance between the right and left gonion points. (B) Bicondylar breadth, defined as the distance between the most lateral points of the mandibular condyles. (C) Bimental breadth, measured as the linear distance between the right and left mental foramina. (D) Symphysis height, measured as the vertical distance from the alveolar crest of the mandibular central incisors to the most inferior point of the mandibular symphysis

CBCT: Cone-beam computed tomography

Statistical Analysis

All analyses were performed using SPSS software (version 26.0; IBM Corp., Armonk, NY). Categorical variables were expressed as frequencies and percentages, and continuous variables as means and standard deviations. The normality of continuous variables was assessed using the Kolmogorov-Smirnov test. A general linear model (GLM) with repeated measures was used to evaluate side (right vs. left) and sex effects on bilateral measurements. Independent samples T tests were used to compare unpaired variables between sexes. Pearson correlation analysis was conducted to assess associations between age and mandibular measurements. Discriminant function analysis was performed to determine which combination of measurements best predicts sex classification.

RESULTS

CBCT images of 199 individuals (72 males and 127 females) were evaluated. The mean age was 32.96 ± 14.66 with minimum 18 and maximum 81 years old individuals. .

Male participants exhibited significantly greater mean values for condylar length ($p=0.035$), minimal ramus breadth ($p=0.001$), mandibular length ($p<0.001$), and coronoid length ($p<0.001$) compared to females (**Table 2**). The effect of side (right vs. left) was not significant for bilateral measurements ($p>0.05$), except for coronoid length ($p=0.017$).

Males demonstrated significantly higher values for bigonial breadth, bicondylar breadth, bimental breadth, and symphyseal height ($p<0.001$), whereas mandibular angles did not differ significantly between sexes ($p>0.05$) (**Table 2**).

	Male, mean $\pm$ SD	Female, mean $\pm$ SD	p
Condylar length	68.76 $\pm$ 3.57	59.27 $\pm$ 2.69	0.035
Coronoid length	63.79 $\pm$ 0.52	55.85 $\pm$ 0.39	<0.001
Mandibular length	77.55 $\pm$ 0.74	73.86 $\pm$ 0.56	<0.001
Minimal ramus breadth	29.95 $\pm$ 0.37	28.36 $\pm$ 0.28	0.001
Mandibular angle	118.79 $\pm$ 0.75	119.33 $\pm$ 0.56	0.566
Bicondylar breadth	113.36 $\pm$ 10.84	104.95 $\pm$ 8.42	<0.001
Bigonial breadth	96.72 $\pm$ 8.42	89.31 $\pm$ 5.32	<0.001
Bimental breadth	46.39 $\pm$ 3.30	43.89 $\pm$ 3.06	<0.001
Symphyseal height	30.76 $\pm$ 4.10	26.92 $\pm$ 3.58	<0.001

SD: Standard deviation

Pearson correlation analysis revealed significant associations between age and right mandibular length ($r=0.204$; $p=0.004$), left mandibular length ($r=0.142$; $p=0.046$), bicondylar breadth ($r=-0.224$; $p=0.002$), and symphyseal height ($r=-0.290$; $p<0.001$).

Discriminant function analysis was conducted to determine whether mandibular measurements could accurately predict sex. The analysis produced one significant canonical discriminant function with a canonical correlation of 0.751 and Wilks' Lambda of 0.436 ($\chi^2(14)=157.58$, $p<0.001$), explaining 100% of the variance.

The most influential variables based on structure matrix coefficients were left coronoid length (0.791), right coronoid length (0.649), left condylar length (0.492), and bigonial breadth (0.477) (**Table 3**). The classification accuracy was 89.9% for original grouped cases, indicating a high level of predictive validity. In the classification matrix, 7 males (9.7%) were misclassified as females, and 10 females (7.9%) were misclassified as males, resulting in an overall accuracy of 91.5%. Cross-validated classification yielded 89.4% correct predictions. In stepwise discriminant analysis, 89.9% of original grouped cases correctly classified with was used left condylar length, left mandibular length, left coronoid length, left mandibular angle, bicondylar breadth, bigonial breadth parameters (**Table 4**).

Table 3. Fisher's linear discriminant functions for sex classification based on mandibular morphometric measurements

Parameter	Function
Coronoid length (left)	0.791
Coronoid length (right)	0.649
Condylar length (left)	0.492
Bigonial breadth	0.477
Symphyseal height	0.433
Bicondylar breadth	0.381
Bimental breadth	0.337
Mandibular length (left)	0.240
Mandibular length (right)	0.232
Minimal ramus breadth (right)	0.221
Minimal ramus breadth (left)	0.181
Condylar length (right)	0.085
Mandibular angle (left)	-0.040
Mandibular angle (right)	-0.028

Table 4. Discriminant analysis of sex using mandibular morphometric parameters

		Predicted group membership		
	Sex	Male [n (%)]	Female [n (%)]	Total, n
Original	Male	65 (90.3)	7 (9.7)	72
	Female	10 (7.9)	117 (92.1)	127
Cross-validated	Male	64 (88.9)	8 (11.1)	72
	Female	13 (10.2)	114 (89.8)	127

The ICC was calculated as 0.986, which reflects excellent intra-observer agreement and further strengthens the validity of the study findings.

DISCUSSION

Sex estimation is a fundamental aspect of forensic identification, especially in cases involving fragmented, incomplete, or severely decomposed remains. The mandible, as the strongest and most durable bone of the craniofacial complex, offers critical morphometric indicators that can be employed in the absence of more complete skeletal structures.¹⁸

In this study, mandibular morphology was evaluated using CBCT in a Turkish adult population. Our findings revealed statistically significant sexual dimorphism in several mandibular parameters, including condylar and coronoid lengths, mandibular length, minimal ramus breadth, and transverse dimensions such as bigonial and bicondylar breadths. In contrast, mandibular angle did not differ significantly between sexes, with females displaying slightly higher values-consistent with previous studies conducted in Turkish populations.^{9,19} The discriminant function analysis yielded a remarkable classification accuracy of 91.5%, with left coronoid length, left condylar length, and bigonial breadth emerging as the most influential predictors. Misclassifications were relatively balanced between sexes (9.7% in males vs. 7.9% in females), indicating that the model did not exhibit sex-related bias. Errors were mostly observed in cases with borderline values of mandibular length and bicondylar breadth, where partial overlap between male and female measurements may have contributed to reduced discriminability.

It is well established that skeletal characteristics vary across populations.⁷ Kharoshah et al.²⁰ reported an 83.9% classification accuracy in an Egyptian sample based on bicondylar breadth, mandibular angle, minimum ramus breadth, and ramus length. In Northeast China, Zheng et al.²¹ reported an overall accuracy of 87.4% using CBCT measurements. Similarly, Dong et al.²² demonstrated predictive accuracies ranging from 80.5% to 84.2% across various multivariate models in a Chinese sample, with bicondylar breadth showing pronounced dimorphism. de Oliveira Gamba et al.²³ achieved 95.1% correct classification in a Brazilian population using parameters such as ramus length, gonion-gnathion length, and bigonial breadth. Alias et al.²⁴ reported an overall classification accuracy of 78.5%, identifying bigonial breadth and condylar height as the strongest discriminators. Conversely, Kano et al.,²⁵ in a Japanese population, did not find significant sexual dimorphism in gnathion-condyle distances or bicondylar breadth. These different results reveal

the impact of population-specific variation and highlight the importance of developing population-specific discriminant models.

The discrepancies among populations regarding the sexual dimorphism of the mandible may be attributed to several factors. Genetic and ethnic variability is an important determinant of craniofacial morphology, leading to population-specific patterns.^{26,27} Environmental and functional factors such as stressful lifestyle and dietary habits can also influence mandible development and measurements.^{28,29} In our sample, the absence of significant dimorphism in the mandibular angle, along with slightly higher values in females, suggests that population-specific skeletal characteristics and age-related remodeling may have contributed to this result.

Several studies in the Turkish population have investigated sex prediction using mandibular and cranial measurements. For example, Okkesim and Erhamza⁷ reported significant sex-based differences in mandibular ramus dimensions using CBCT. Inci et al.¹⁹ achieved high accuracy rates (51-95.6%) for sex estimation based on mandibular ramus measurements, with superior ramus height being the most discriminative parameter. Meyvaci et al.³⁰ evaluated angular mandibular parameters and reported a 71.5% accuracy rate using a discriminant function based on four angular measurements. These findings confirm the presence of sexual dimorphism in both mandibular and cranial structures among the Turkish population. However, prediction accuracy depends on the selected parameters and methodology.

In a recent meta-analysis, mandibular measurements were found to be valuable indicators for sex estimation, as males typically exhibit greater bicondylar and bigonial breadths compared to females.¹⁸ In our study, coronoid length, left condylar length, and bigonial breadth were found to be most distinctive parameters. Bicondylar and bigonial breadth were found to be sexually discriminative parameters align with the meta-analysis results. They reported a heterogeneity in the results of studies regarding the presence of sexual dimorphism in mandibular angle measurements, interestingly they reported sexual dimorphism on CBCT scans, this distinction is not evident in computed tomography scans. In our study mandibular angle was not different between males and females on CBCT scans and females had slightly higher values. While some studies have reported higher values in females,^{19,21} others suggest that it lacks consistent dimorphic patterns across populations.^{20,23} Bulut et al.³¹ found that the mandibular angle was only significantly different in older age groups (60-80 years), suggesting that age-related changes may confound its forensic utility.

In addition to sex-related differences, age was significantly correlated with certain mandibular parameters, including mandibular length, bicondylar breadth, and symphyseal height. The mandible undergoes morphological alterations in size and structure throughout human growth. Studies have demonstrated a strong association between chronological age and mandibular morphology, particularly in the ramus region.³² In a study investigating the direction of craniofacial growth from late adolescence to late adulthood, it was concluded that mandibular growth direction differed between

genders, with males exhibiting more forward rotation; this may explain why the mandibular angle tends to be lower in males.³³ In our study, although the mandibular angle showed slightly higher values in females, it did not demonstrate significant sexual dimorphism, which may partly reflect its age dependency. Therefore, while CBCT-derived mandibular measurements are highly useful for sex estimation, parameters that are susceptible to age-related changes should be interpreted with caution in forensic practice.

The integration of advanced medical imaging techniques such as CBCT into forensic science has significantly enhanced the accuracy and reliability of skeletal assessments. Compared to conventional cephalometric radiographs, CBCT offers superior resolution, three-dimensional visualization, and non-invasive measurement capabilities, reducing inter-observer variability and increasing objectivity.^{5,8} These advantages are particularly beneficial when evaluating sexually dimorphic features of the mandible.

Recent advances in AI-based approaches have substantially enhanced medical imaging, with notable implications for forensic science, including tasks such as age estimation and sex determination. In forensic contexts, accurate prediction of age and gender is essential, and deep learning offers the potential to increase both reliability and efficiency in these assessments. Recent studies support the potential of automated pipelines for forensic sex estimation. Küchler et al.¹³ demonstrated that combining mandibular and dental measurements with machine-learning algorithms achieved high predictive performance for sex determination. Similarly, a study evaluated CBCT-derived morphometric features across multiple classifiers, confirming their accuracy for sex identification.¹² More recently, Pishghadam et al.¹⁵ introduced an explainable multi-task deep learning framework applied to dental CBCT images, enabling simultaneous age estimation and sex classification with attention-based interpretability. These advances highlight that integrating AI based approaches can provide reproducible, observer-independent methods that enhance the forensic applicability of mandibular morphometry.

This study possesses several notable strengths that enhance its scientific validity and potential applicability. First, the relatively large sample size (n=199) contributes to the statistical power and increases the reliability of the conclusions drawn. Second, a comprehensive set of nine bilateral mandibular morphometric parameters was assessed, allowing for a multidimensional evaluation of sexual dimorphism. The use of CBCT imaging—a high-resolution, low-radiation, and non-destructive modality—offered superior anatomical detail and measurement accuracy compared to conventional radiographic techniques. Finally, the application of a broad range of statistical analyses, including GLM, t-tests, Pearson correlation, and both standard and stepwise discriminant function analysis, enriched the robustness of the findings and allowed for the development of predictive models with high classification accuracy.

The results of this study have important implications for both clinical and forensic practices. From a forensic perspective, the high accuracy of sex estimation based on CBCT-

derived mandibular measurements offers a valuable tool for identification in mass disasters, criminal investigations, or culturally sensitive contexts where autopsies are restricted. The mandible's resilience to trauma and postmortem degradation further supports its utility as a key element in forensic anthropology. Clinically, these findings may assist in designing sex-specific diagnostic, surgical, or prosthetic interventions, particularly in maxillofacial and orthodontic fields. Furthermore, the measurement protocol and discriminant function outputs of this study could serve as foundational datasets for developing automated sex estimation algorithms based on machine learning, thereby reducing operator dependency and increasing efficiency.

Future research should aim to include additional cranial and facial measurements, explore sex estimation in subadults, and validate findings in larger, multiethnic samples. Integrating artificial intelligence models—such as CNNs for direct image-based classification, U-Net variants for automated mandibular segmentation, and machine-learning classifiers (e.g., support vector machines, random forests) trained on CBCT-derived measurements—with large annotated datasets may pave the way for fully automated and standardized forensic identification systems.

Limitations

Despite its contributions, the study has several limitations. The analysis was limited to mandibular structures, excluding other potentially informative craniofacial or skeletal landmarks that may improve classification accuracy. Moreover, the retrospective study design may introduce inherent biases related to patient selection and data quality. Although the sample encompassed a wide age range, subgroup analyses by specific age cohorts were not performed. This limits the ability to comprehensively assess age-related variations in mandibular morphology, despite the observed correlations suggesting that age may act as a contributing factor. Since the primary aim of the study was to investigate sex differentiation, detailed age-stratified analyses were beyond the intended scope but should be considered in future research. Finally, the study was conducted on a single Turkish population, which may limit the generalizability of the findings to other ethnic or demographic groups.

CONCLUSION

This study demonstrated that mandibular morphometric parameters obtained from CBCT images exhibit significant sexual dimorphism and can be used effectively for sex estimation in a Turkish population. Among the evaluated measurements, coronoid and condylar lengths, mandibular length, and transverse dimensions such as bigonial and bicondylar breadths showed the highest discriminative power. The discriminant function analysis yielded a high classification accuracy, highlighting the potential of these parameters as reliable forensic indicators. Given the mandible's durability and resistance to postmortem changes, CBCT-based morphometric analysis provides a reproducible and non-invasive approach for personal identification. Future studies incorporating larger and more diverse populations, along with AI approaches—such as machine learning and

deep learning models-are encouraged to further validate and refine these findings for broader forensic and anthropological applications.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Süleyman Demirel University Health Sciences Ethics Committee (Date: 04.07.2025, Decision No: 35).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Sujatha S, Azmi SR, Devi B, Shwetha V, Kumar TP. CBCT-the newfangled in forensic radiology. *J Dent Orofac Res*. 2017;13(2):47-55.
- Akhlaghi M, Vasigh S, Khalighi Z, Yousefinejad V. The value of mandible measurements in gender prediction for the Iranian adult population. *Aust J Forensic Sci*. 2014;46(2):127-135. doi:10.1080/00450618.2013.807358
- Gurushanthappa V, Rajashekarappa MGS. Determination of Gender by the Anthropometric Measurement of Human Mandible Using Ramus Breadth and Mandibular Angle: A Cross Sectional Study from South India. *Med Sci*. 2013;1(2):28-32. doi:10.29387/ms.2013.1.2.28-32
- Kallalli BN, Rawson K, Ramaswamy VK, Zakarneh WH, Singh A, Zingade J. Sex determination of human mandible using metrical parameters by computed tomography: A prospective radiographic short study. *J Indian Acad Oral Med Radiol*. 2016;28(1):7-10. doi:10.4103/0972-1363.189990
- Assari A, Alasmari B, Aleid M, Salem M. Characteristics of mandibular parameters in different age groups a CBCT assessment. *EC Dent Sci*. 2017;14:95-103.
- Gillet C, Costa-Mendes L, Rérolle C, Telmon N, Maret D, Savall F. Sex estimation in the cranium and mandible: a multislice computed tomography (MSCT) study using anthropometric and geometric morphometry methods. *Int J Legal Med*. 2020;134:823-832. doi: 10.1007/s00414-019-02203-0
- Okkesim A, Erhamza TS. Assessment of mandibular ramus for sex determination: retrospective study. *J Oral Biol Craniofac Res*. 2020;10(4): 569-572. doi: 10.1016/j.jobcr.2020.07.019
- Albalawi AS, Alam MK, Vundavalli S, Ganji KK, Patil S. Mandible: an indicator for sex determination - a three-dimensional cone-beam computed tomography study. *Contemp Clin Dent*. 2019;10(1):69-73. doi: 10.4103/ccd.ccd_313_18
- İlgüy D, İlgüy M, Ersan N, Dölekoğlu S, Fişekçioğlu E. Measurements of the foramen magnum and mandible in relation to sex using CBCT. *J Forensic Sci*. 2014;59(3):601-605. doi:10.1111/1556-4029.12376
- Rames JD, Hussein SM, Shehab AA, et al. Mandibular gender dimorphism: the utility of Artificial Intelligence and statistical shape modeling in skeletal facial analysis. *Aesthet Plast Surg*. 2024;48(21):4272-4279. doi: 10.1007/s00266-024-04300-x
- Suzuki K, Nakano H, Inoue K, et al. Examination of new parameters for sex determination of mandible using Japanese computer tomography data. *Dentomaxillofac Radiol*. 2020;49(5):20190282. doi:10.1259/dmfr.20190282
- Baban MTA, Mohammad DN. The accuracy of sex identification using CBCT morphometric measurements of the mandible, with different machine-learning algorithms-a retrospective study. *Diagnostics*. 2023;13(14):2342. doi: 10.3390/diagnostics13142342
- Küchler EC, Kirschneck C, Marañón-Vásquez GA, et al. Mandibular and dental measurements for sex determination using machine learning. *Sci Rep*. 2024;14(1):9587. doi:10.1038/s41598-024-59556-9
- Vila-Blanco N, Varas-Quintana P, Aneiros-Ardao Á, Tomás I, Carreira MJ. Automated description of the mandible shape by deep learning. *Int J Comput Assist Radiol Surg*. 2021;16(12):2215-2224. doi:10.1007/s11548-021-02474-2
- Pishghadam N, Esmailyfard R, Paknahad M. Explainable deep learning for age and gender estimation in dental CBCT scans using attention mechanisms and multi task learning. *Sci Rep*. 2025;15(1):18070. doi:10.1038/s41598-025-03305-z
- Kumar MP, Lokanadham S. Sex determination & morphometric parameters of human mandible. *Int J Res Med Sci*. 2013;1(2):93-96. doi: 10.5455/2320-6012.ijrms20130511
- Direk F, Uysal II, Kivrak AS, Dogan NU, Fazliogullari Z, Karabulut AK. Reevaluation of mandibular morphometry according to age, gender, and side. *J Craniofac Surg*. 2018;29(4):1054-1059. doi: 10.1097/SCS.0000000000004293
- Munhoz L, Okada S, Hisatomi M, Yanagi Y, Arita ES, Asaumi J. Are computed tomography images of the mandible useful in age and sex determination? A forensic science meta-analysis. *J Forensic Odontostomatol*. 2024;42(1):38-57. doi:10.5281/zenodo.11058169
- Inci E, Ekizoglu O, Turkay R, et al. Virtual assessment of sex: linear and angular traits of the mandibular ramus using three-dimensional computed tomography. *J Craniofac Surg*. 2016;27(7):e627-e632. doi: 10.1097/SCS.0000000000002979
- Kharoshah MAA, Almadani O, Ghaleb SS, Zaki MK, Fattah YAA. Sexual dimorphism of the mandible in a modern Egyptian population. *J Forensic Leg Med*. 2010;17(4):213-215. doi:10.1016/j.jflm.2010.02.005
- Zheng J, Ni S, Wang Y, Zhang B, Teng Y, Jiang S. Sex determination of Han adults in Northeast China using cone beam computer tomography. *Forensic Sci Int*. 2018;289:450. doi:10.1016/j.forsciint.2018.05.036
- Dong H, Deng M, Wang W, Zhang J, Mu J, Zhu G. Sexual dimorphism of the mandible in a contemporary Chinese Han population. *Forensic Sci Int*. 2015;255:9-15. doi:10.1016/j.forsciint.2015.06.010
- de Oliveira Gamba T, Alves MC, Haiter-Neto F. Mandibular sexual dimorphism analysis in CBCT scans. *J Forensic Leg Med*. 2016;38:106-110. doi: 10.1016/j.jflm.2015.11.024
- Alias A, Ibrahim A, Abu Bakar SN, et al. Anthropometric analysis of mandible: an important step for sex determination. *Clin Ter*. 2018;169(5): e217-e223. doi:10.7417/CT.2018.2082
- Kano T, Oritani S, Michiue T, et al. Postmortem CT morphometry with a proposal of novel parameters for sex discrimination of the mandible using Japanese adult data. *Leg Med (Tokyo)*. 2015;17(3):167-171. doi:10.1016/j.legalmed.2014.12.009
- Ichijo Y, Takahashi Y, Tsuchiya M, et al. Relationship between morphological characteristics of hyoid bone and mandible in Japanese cadavers using three-dimensional computed tomography. *Anat Sci Int*. 2016;91(4):371-381. doi:10.1007/s12565-015-0312-z
- Silva LP, Leite RB, Sobral AP, et al. Oral and maxillofacial lesions diagnosed in older people of a Brazilian population: a multicentric study. *J Am Geriatr Soc*. 2017;65(7):1586-1590. doi:10.1111/jgs.14815
- Holmes MA, Ruff CB. Dietary effects on development of the human mandibular corpus. *Am J Phys Anthropol*. 2011;145(4):615-628. doi:10.1002/ajpa.21554
- Filho J, Manzi F, de Freitas D, Bóscolo F, de Almeida S. Evaluation of temporomandibular joint in stress-free patients. *Dentomaxillofacial Radiol*. 2007;36(6):336-340. doi:10.1259/dmfr/17973079

30. Meyvacı SS, Bulut DG, Öztürk AT, Ankaralı H. Gender estimation from angular parameters of mandible in Turkish adults. *Med J West Black Sea*. 2021;5(2):240-247. doi:10.29058/mjwbs.873771
31. Bulut O, Freudenstein N, Hekimoglu B, Gurcan S. Dilemma of gonial angle in sex determination: sexually dimorphic or not? *Am J Forensic Med Pathol*. 2019;40(4):361-365. doi:10.1097/PAF.0000000000000500
32. Motawei SM, Helaly AM, Aboelmaaty WM, Elmahdy K, Shabka OA, Liu H. Length of the ramus of the mandible as an indicator of chronological age and sex: a study in a group of Egyptians. *Forensic Sci Int Rep*. 2020;2:100066. doi:10.1016/j.fsir.2020.100066
33. Pecora NG, Baccetti T, McNamara Jr JA. The aging craniofacial complex: a longitudinal cephalometric study from late adolescence to late adulthood. *Am J Orthod Dentofacial Orthop*. 2008;134(4):496-505. doi:10.1016/j.ajodo.2006.11.022

Bicycle accidents and injury patterns in emergency department admissions: a single-center experience

 Tamer Durdu,  Yıldız Yıldırım Çelik,  Nihal Ertürk,  Habibe Selmin Özensoy,
 Furkan Altaş,  Hakan Oğuztürk

Department of Emergency Medicine, Ankara Bilkent City Hospital, Ankara, Türkiye

Cite this article as: Durdu T, Yıldırım Çelik Y, Ertürk N, Özensoy HS, Altaş F, Oğuztürk H. Bicycle accidents and injury patterns in emergency department admissions: a single-center experience. *J Health Sci Med.* 2025;8(6):1018-1022.

Received: 23.08.2025

Accepted: 22.09.2025

Published: 25.10.2025

ABSTRACT

Aims: The aim of this study was to define injury patterns of bicycle accidents in order to guide preventive strategies and resource allocation in emergency departments. Bicycle-related injuries are a significant and growing public health concern. In Türkiye, the increasing popularity of bicycles for transportation and recreation has been paralleled by rising emergency department (ED) admissions. Understanding the epidemiology, mechanisms, and outcomes of these injuries is crucial for developing preventive strategies.

Methods: This retrospective single-center study included patients admitted to the Emergency Department of Ankara Bilkent City Hospital between February 2019 and December 2024 due to bicycle-related injuries. Data were obtained from the hospital information system and included demographics, injury mechanism, type of bicycle, injury site, consultations, interventions, length of stay, and outcomes. Patients with incomplete records (n=43) were excluded. Statistical analyses were performed using SPSS version 25, with $p < 0.05$ considered significant.

Results: A total of 317 patients were analyzed, 269 (84.9%) were male and 48 (15.1%) were female. Most patients were Turkish (92.1%). The number of cases increased over the years, peaking in 2024. Helmet use was documented in 134 children (59.8%) and 76 adults (81.7), and its absence was associated with higher rates of head and facial injuries ($p < 0.05$). The most frequent accident mechanism was falling from the bicycle (68.3%), followed by collisions with stationary objects (14.1%). Urban/commuter bicycles (49.3%) were the most common type, while electric bicycles accounted for 19.6%. The most common injury sites were upper extremities (76.0%), face (40.2%), and head (35.2%). Abdominal injuries occurred in 7.9% of cases, with splenic injury being the most frequent (4.4%). No mortality was observed during the study period.

Conclusion: Bicycle-related injuries predominantly affect young males and frequently involve extremities, head, and face. Despite moderate helmet use, craniofacial injuries remain common. Abdominal trauma, though less frequent, represents a serious clinical concern, particularly in pediatric cases. These findings highlight the need for improved safety measures, widespread helmet use, and public education to reduce the burden of bicycle-related trauma.

Keywords: Bicycle, injuries, emergency department

INTRODUCTION

Bicycles are widely used worldwide as a means of transportation, sport, and recreation, but they also carry a significant risk of trauma-related injuries. Interest in cycling is increasing in our country, and in parallel, emergency room visits due to injuries related to bicycle accidents are increasing.¹

These injuries most often occur as a result of falls from bicycles and striking fixed or moving objects. Fatal and serious injuries often occur in collisions with motor vehicles.² Within the pediatric age group, children aged 10-15 have the highest mortality rate.³ Fractures are the most common cause of morbidity, and traumatic brain injuries are the most common cause of mortality and long-term disability.⁴ Handlebar injuries to the abdomen and pelvis are also a

significant cause of morbidity.⁵ Passengers and children carried in bicycle-mounted seats are at risk, particularly if their feet, legs, or clothing become entangled in the spokes and chain.⁶

The aim of this study was to define injury patterns of bicycle accidents in order to guide preventive strategies and resource allocation in emergency departments. Bicycle-related injuries are a significant and growing public health concern. In Türkiye, the increasing popularity of bicycles for transportation and recreation has been paralleled by rising emergency department (ED) admissions. Understanding the epidemiology, mechanisms, and outcomes of these injuries is crucial for developing preventive strategies.

Corresponding Author: Tamer Durdu, tamdurdu@hotmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

METHODS

This is a single-center retrospective study conducted at the Ankara Bilkent City Hospital Emergency Department between February 2019 and December 2024. This hospital is the largest hospital in Ankara and Türkiye, and is also a trauma center. Data were obtained from the hospital's automation system and include all trauma-related injuries, treatments, hospitalizations, and outcomes. The study was conducted after receiving approval from the Ankara Bilkent City Hospital Medical Researches Ethics Committee (Date: 28.05.2025, Decision No: TABED 2/1211/2025). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

All data were recorded by trained healthcare personnel in the emergency department. Records included patient demographics, date and time of injury, location, mechanism of injury, emergency room, interventions performed, consultations requested, site of injury, diagnosis of injury, hospitalization area, day of hospitalization, 30-day mortality, and discharge status. The number of patients included in our study was 317. Forty-three patients with missing data were excluded.

Statistical Analysis

Data were obtained retrospectively from hospital records. Statistical analyses were performed using IBM SPSS Statistics (version 25).

Distributional characteristics of continuous variables were assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests, histograms, and Q-Q plots. Normally distributed variables were summarized as mean±standard deviation, and non-normally distributed variables were summarized as median (minimum-maximum).

Continuous variables between two independent groups were compared using the independent samples T test. In cases where distribution assumptions were not met, the Mann-Whitney U test was used.

Categorical variables were expressed as frequencies and percentages. The Chi-square test (Pearson Chi-square) was used for comparisons between groups, and the Fisher's exact test was used when the expected cell counts were low. A p value of <0.05 was considered statistically significant in all tests.

RESULTS

A total of 317 patients were included, of whom 269 (84.9%) were male and 48 (15.1%) were female. The majority were aged ≤18 years (n=224, 70.7%), while 93 (29.3%) were adults. For patients younger than 18 years, the mean age was 11.26 years (SD=0.23; range: 3-17 years), with a median age of 11 years (IQR=5). For patients aged 18 years and above, the mean age was 33.09 years (SD=11.924; range: 18-69 years), with a median age of 29.5 years (IQR=8-159). Most patients were Turkish (92.1%), with smaller proportions of Iraqi, Syrian, Afghan, Azerbaijani, American, and Canadian nationals. The number of cases increased progressively across the study years, ranging from 22 (6.9%) in 2019 to 86 (27.1%) in 2024 (Figure 1). Seasonal variation was evident, with the peak in July (21.5%) and the lowest number in December (0.9%) (Figure 2).

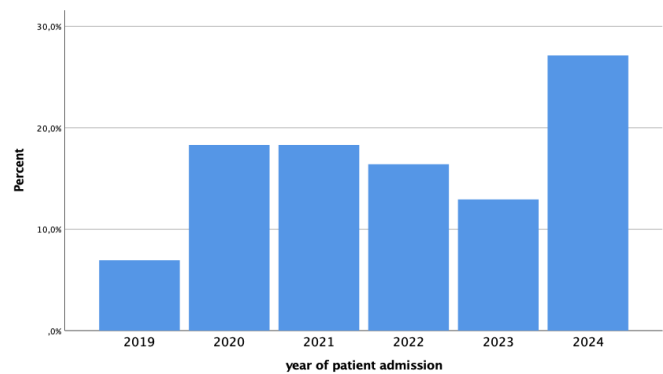


Figure 1. Year of patient admission

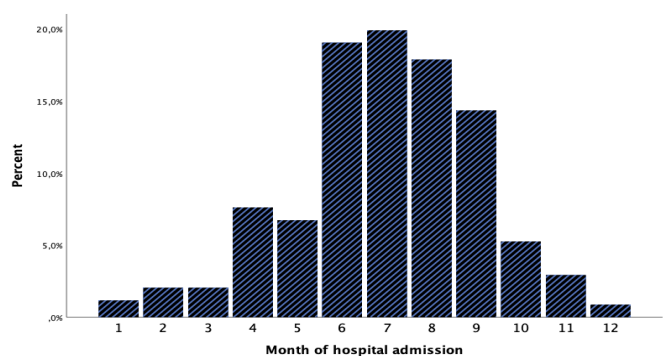


Figure 2. Percentage of hospital admissions by month

Helmet use was documented in 134 children (59.8%) and 76 adults (81.7%). A total of 9 patients (2.8%) were bicycle couriers, 8 of whom were adults. Regarding bicycle type, city bikes (54%) predominated among children, followed by children's training bicycles (20.5%), mountain bikes (17.4%), electric bikes (6.3%), and racing bikes (1.8%) (Figure 3). Among adults, electric bikes were most common (49.5%), followed by city bikes (38.7%), racing bikes (5.4%), and mountain bikes (6.5%) (Table 1).

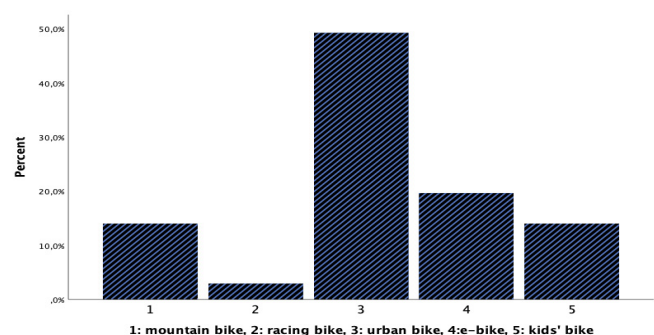


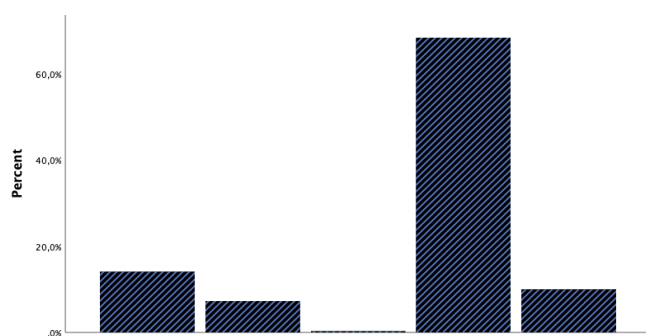
Figure 3. Percentage of bicycle types

Falls from bicycles were the leading mechanism of injury (68.3% in children, 78.5% in adults). Collisions with stationary objects, other bicycles, or motor vehicles were less frequent (Figure 4). Most incidents occurred on streets and avenues (68.3% in children, 51.6% in adults), with smaller proportions in parks, main roads, bicycle paths, and off-road areas.

In terms of triage, children were most frequently admitted to the yellow zone (54.5%), followed by the green (23.7%) and red zones (21.9%). A similar distribution was observed in adults (53.8%, 24.7%, and 21.5%, respectively).

Table 1. Demographics and event characteristics

Variable	n (%)
Total patients	317 (100)
Male	269 (84.9)
Female	48 (15.1)
≤18 years	224 (70.7)
>18 years	93 (29.3)
Nationality: Turkish	292 (92.1)
Nationality: Other	25 (7.9)
Helmet use (≤18 years)	134 (59.8)
No helmet (≤18 years)	90 (40.2)
Helmet use (>18 years)	76 (81.7)
No helmet (>18 years)	17 (18.3)
Courier occupation (total)	9 (2.8)
Most common bicycle type (≤18 years)	City bikes 121 (54.0)
Most common bicycle type (>18 years)	Electric bikes 46 (49.5)
Most common accident type (≤18 years)	Falls 143 (68.3)
Most common accident type (>18 years)	Falls 73 (78.5)
Most common accident location (≤18 years)	Street/avenues 153 (68.3)
Most common accident location (>18 years)	Street/avenues 48 (51.6)

**Figure 4. Percentage of accident types**

Injury distribution differed by age. In children, the upper extremities (71.4%) and lower extremities (62.1%) were most commonly affected, followed by the face (43.3%), head (37.5%), and chest (15.6%). In adults, lower extremity injuries predominated (81.7%), with additional involvement of the upper extremities (28.6%), face (32.3%), head (29%), and pelvis (22.6%) (Table 2).

Table 2. Injury patterns and outcomes

Variable	n (%) / mean (SD)
Most common injuries (≤18 years)	Dermoabrasion/laceration 212 (94.6), long bone fractures 55 (24.6), maxillofacial 44 (19.6)
Most common injuries (>18 years)	Dermoabrasion/laceration 102 (98.1), long bone fractures 28 (26.9), maxillofacial 11 (10.6)
Consultations requested	Orthopedics 118 (37.2), plastic surgery 57 (18.0), others <10% each
Hospital stay (overall)	Mean 1.55 (SD 0.41), median 0, range 0-84
Hospital stay (>18 years)	Mean 1.75 (SD 0.41), median 1, range 0-25
ED observation	114 (35.2)
Ward admission	28 (8.5)
ICU + ward	17 (5.9)
Ongoing ICU	7 (2.1)
Discharged	151 (48.4)
Mortality	0 (0%)

SD: Standard deviation, ED: Emergency department, ICU: Intensive care unit

Accidents occurred most frequently in the evening hours (36.9%) between 18:01-24:00 and afternoon (36.6%) between 12:01-18:00.

The most frequent final diagnoses included dermoabrasion/laceration (92.1%), long bone fractures (26.1%), and maxillofacial trauma (17.3%). Less frequent injuries comprised intracranial hemorrhage (3.1%), splenic injury (4.7%), gastrointestinal injury (2.2%), rib fracture (0.6%), and liver injury (0.9%). Orthopedics (37.2%) and plastic surgery (18%) were the most frequently consulted specialties, while more than half of the patients (50.7%) required no specialty consultation. Thoracic injuries were generally minor and managed conservatively.

The mean length of hospital stay among 317 patients was 1.55 days (SD=5.365; range: 0-84 days; variance=28.784). Given the skewed distribution of length of hospital stay, the median length of stay was 0 days (IQR=0-1 days), indicating that the majority of patients were treated as outpatients. Overall, 114 patients (35.2%) were observed in the emergency department, 28 (8.5%) were admitted to the ward, 17 (5.9%) required combined ward and intensive care management, 7 (2.1%) remained in intensive care, and 151 (48.4%) were discharged after initial treatment.

Injury pattern analysis showed that dermoabrasion/laceration (94.6%), long bone fractures (24.6%), and maxillofacial trauma (19.6%) were most common among children. In adults, dermoabrasion/laceration (98.1%) and long bone fractures (26.9%) predominated, followed by maxillofacial trauma (10.6%) and intracranial hemorrhage (5.8%). No deaths occurred during the study period (Table 3).

Table 3. Comparison of injury sites between patients ≤18 and >18 years of age

Injury site	≤18 years (n, %)	>18 years (n, %)	p-value
Head injury present	84 (37.5%)	27 (29%)	p=0.094
Head injury absent	140 (62.5%)	66 (71%)	
Facial injury present	97 (43.3%)	30 (32.3%)	p=0.044
Facial injury absent	127 (56.7%)	63 (67.7%)	
Eye injury present	27 (12.1%)	1 (1.1%)	p=0.0001
Eye injury absent	197 (87.9%)	92 (98.9%)	
Dental injury present	30 (13.4%)	2 (2.2%)	p=0.001
Dental injury absent	194 (86.6%)	91 (97.8%)	
Thoracic injury present	35 (15.6%)	11 (11.8%)	p=0.245
Thoracic injury absent	189 (84.4%)	82 (88.2%)	

Statistical test: Pearson Chi-square

DISCUSSION

In this single-center cohort of 317 ED admissions for bicycle-related trauma, patients were predominantly male and pediatric, with clear summertime and evening peaks. These distributions mirror population-level patterns of bicycling exposure and ED use reported in national and international series, wherein cycling-related traumatic brain injury and ED visits are more frequent among males and younger riders.^{1,3} Our year-on-year rise and strong July peak likely reflect growing bicycle use locally as well as seasonal ridership trends.^{1,2}

Falls comprised the leading mechanism in both age groups, whereas motor-vehicle collisions were comparatively infrequent. This aligns with literature showing that while collisions with motor vehicles drive the highest severity and mortality burden, simple falls dominate ED presentations, particularly among younger or less experienced riders.^{2-4,8} The predominance of street/avenue locations in our series is likewise consistent with urban exposure patterns.²

Despite a relatively high reported helmet use overall, head and facial injuries remained common, and non-helmeted patients had significantly higher rates of head/facial trauma in unadjusted comparisons. These findings accord with meta-analytic evidence that helmets reduce head injury risk by roughly 40-60% and facial injury risk to a lesser extent.¹⁰ Residual head/facial injury in helmeted riders is biologically plausible given suboptimal fit, limited facial coverage of conventional helmets, and the role of angular acceleration in traumatic brain injury.^{4,10} From a prevention perspective, our data support sustained promotion of helmet use, fit/retention education, and consideration of face-protective designs in higher-risk subgroups.^{4,10}

Upper and lower extremity injuries were the most frequent regional diagnoses, consistent with forward/sideways falls described in prior series and reviews.^{3,4} The high demand for orthopedic consultation (37.2%) in our cohort reflects this burden and underscores planning needs for ED radiography, immobilization, and fracture care pathways.^{4,9} Maxillofacial trauma was also common, reinforcing the value of coordinated ED-dental/ENT-plastic surgery triage for laceration repair, dental avulsion, and orbit/zygomatic injuries.⁴

Adults most frequently rode electric bicycles in our sample. Prior studies indicate distinctive injury patterns and potentially greater severity among e-bike users, plausibly linked to higher speeds and rider mix.⁹ While our study was not powered to compare e-bike versus conventional bikes formally, the case mix suggests targeted messaging on speed control, braking technique, and infrastructure that accommodates e-bike acceleration profiles.^{4,9}

Although less common overall, abdominal injuries occurred in ~8% of patients, with splenic trauma most frequent—particularly in children. This conforms to pediatric blunt abdominal trauma literature and to the well-recognized “handlebar injury” mechanism, which can produce splenic, pancreatic, and hollow-viscus injuries and is prone to delayed diagnosis.¹¹⁻¹⁴ Our zero mortality and largely conservative management are compatible with contemporary pediatric protocols, yet these injuries demand vigilance: meticulous serial exams and liberal use of ultrasound/CT in compatible mechanisms remain essential.^{11,12,14-15} The World Society of Emergency Surgery guidance provides a useful framework for evaluating possible bowel injury when clinical trajectory or imaging is equivocal.¹⁵

Emergency Department observation and same-day discharge were common, length of stay was short, and no deaths occurred—features typical of fall-dominant cohorts with low rates of high-energy collision.^{1,3,4} These patterns support resource allocation toward rapid imaging, laceration and

fracture management, and streamlined specialist access (orthopedics, plastics/dental/ENT, ophthalmology) while maintaining clear escalation pathways for suspected intracranial or intra-abdominal injury.^{4,8,11-15}

Public health and prevention. Our results reinforce three tractable opportunities: (i) sustained, age-tailored promotion of helmet use and correct fit;¹⁰ (ii) rider education on fall-prevention skills and e-bike speed management;^{4,9} and (iii) child-focused messaging on handlebar risks, including protective equipment and safe riding environments.¹¹⁻¹⁴ Given the rising caseload, these strategies—complemented by infrastructure that separates cyclists from high-speed traffic—may reduce ED demand and morbidity.^{1,2,4,10}

Limitations

The retrospective, single-center design may limit generalizability, and unmeasured confounding (e.g., exposure time, riding context) cannot be excluded. Importantly, formal trauma severity scores (e.g., ISS) were unavailable, precluding standardized severity adjustment—an acknowledged limitation that future prospective, multicenter studies should address.

CONCLUSION

This study provides a comprehensive evaluation of bicycle-related injuries presenting to the emergency department of a major trauma center. The findings demonstrate that most injuries result from relatively minor mechanisms such as falls, with upper extremity, facial, and head trauma being the most common. Helmet use was relatively high in our cohort, yet head and facial injuries remained frequent, underscoring the need for broader prevention strategies and increased public awareness. Although abdominal injuries were less common, they carried the potential for serious complications, particularly in children. Importantly, no mortality was observed, reflecting the predominance of relatively minor trauma and the effectiveness of timely management. Preventive strategies such as promoting helmet use, increasing awareness of electric bicycle-related risks, and improving infrastructure for safer cycling remain crucial. Further multicenter prospective studies are needed to validate these findings and to guide evidence-based strategies aimed at reducing the burden of bicycle-related injuries. This study’s strengths include the large sample size, comprehensive six-year timeframe, and detailed analysis of both pediatric and adult patients, which provide valuable insights into bicycle-related trauma.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Ankara Bilkent City Hospital Medical Researches Ethics Committee (Date: 28.05.2025, Decision No: TABED 2/1211/2025).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Sarmiento K, Haileyesus T, Waltzman D, Daugherty J. Emergency department visits for bicycle-related traumatic brain injuries among children and adults-United States, 2009–2018. *MMWR Morb Mortal Wkly Rep*. 2021;70(19):693–697. doi:10.15585/mmwr.mm7019a1
2. Goel R. Population-level estimate of bicycle use and fatality risk in a data-poor setting. *Int J Inj Control Saf Promot*. 2023;30(3):333–337. doi:10.1080/17457300.2023.2172737
3. Teisch LF, Allen CJ, Tashiro J, et al. Injury patterns and outcomes following pediatric bicycle accidents. *Pediatr Surg Int*. 2015;31(11):1021–1025. doi:10.1007/s00383-015-3756-2. PMID:26269203
4. Hyman MJ, Charney C, Arends RW, Heffernan TE, Lawson LA, Lee ES. Trauma management of bicycle injuries. *Curr Trauma Rep*. 2024. doi:10.1007/s40719-024-00335-4
5. Biyyam DR, Hwang S, Patel MC, Bardo DME, Bailey SS, Youssfi M. CT findings of pediatric handlebar injuries. *RadioGraphics*. 2020;40(3):815–826. doi:10.1148/rg.2020190126
6. Agarwal A, Pruthi M. Bicycle-spoke injuries of the foot in children. *J Orthop Surg (Hong Kong)*. 2010;18(3):338–341. doi:10.1177/230949901001800316
7. Reuvers R, van Beeck EF, van Dijck S, et al. Cost-effectiveness of mandatory helmet use. *BMC Public Health*. 2020;20:251. doi:10.1186/s12889-020-08544-5
8. Park JC, Jang EJ, Shin SD, et al. Epidemiology and risk factors for bicycle-related severe head injury: a single-center study of the emergency department patients. *BMC Pediatr*. 2017;17(1):27. doi:10.1186/s12887-017-0782-7
9. Tenenbaum S, Weltsch D, Bariteau JT, et al. Orthopaedic injuries among electric bicycle users. *Injury*. 2017;48(10):2140–2144. doi:10.1016/j.injury.2017.07.012
10. Olivier J, Creighton P. Bicycle injuries and helmet use: a systematic review and meta-analysis. *Int J Epidemiol*. 2017;46(1):278–292. doi:10.1093/ije/dyw153
11. Bašković M, Grgić O, Medved I, et al. The diagnosis and management of pediatric blunt abdominal trauma. *Children (Basel)*. 2024;11(7):829. doi:10.3390/children11070829
12. Cheung R, Shukla M, Akers KG, Farooqi A, Sethuraman U. Bicycle handlebar injuries: systematic review. *Am J Emerg Med*. 2022;51:13–21. doi:10.1016/j.ajem.2021.09.043
13. Bašković M, Dumančić J, Medved I, et al. The diagnosis and management of pediatric blunt abdominal trauma. *Diagnostics (Basel)*. 2024;14(20):2257. doi:10.3390/diagnostics14202257
14. Nadler EP, Potoka DA, Shultz BL, Morrison KE, Ford HR, Gaines BA. The high morbidity associated with handlebar injuries in children. *J Trauma*. 2005;58(6):1171–1174. doi:10.1097/01.TA.0000170107.21534.7A
15. Smyth L, Gaarder C, Naidoo N, et al. WSES guidelines on blunt and penetrating bowel injury: diagnosis, investigations, and treatment. *World J Emerg Surg*. 2022;17:13. doi:10.1186/s13017-022-00418-y

Impact of atrial fibrillation on cognitive decline and cortical atrophy in Alzheimer's disease

Özlem Totuk, Hasan Can Güdek, Şevki Şahin

Department of Neurology, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, İstanbul, Türkiye

Cite this article as: Totuk Ö, Güdek HC, Şahin Ş. Impact of atrial fibrillation on cognitive decline and cortical atrophy in Alzheimer's disease. *J Health Sci Med.* 2025;8(6):1023-1029.

Received: 24.08.2025

Accepted: 23.09.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to evaluate the impact of atrial fibrillation (AF) on cognitive function and neuroradiological markers in patients with Alzheimer's disease (AD).

Methods: A total of 800 patients were screened at the Dementia Outpatient Clinic of Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital. Thirty-three AD patients with confirmed AF were included, alongside 31 age-, sex, and education-matched AD patients without AF as controls. Cognitive function was assessed using the Mini-Mental State Examination (MMSE), while mood symptoms were evaluated with the Geriatric Anxiety Scale and the Geriatric Depression Scale. Brain MRI scans were analyzed using Fazekas, Medial Temporal Atrophy (MTA), and Koedam scores. Due to the modest sample size, this study is best interpreted as exploratory in nature.

Results: After six months, the AF-positive group exhibited a significant decline in MMSE scores, whereas the AF-negative group showed no significant change. Koedam scores were significantly higher in the AF-positive group, indicating more advanced cortical atrophy. Regression analysis revealed that AF had a borderline significant negative impact on cognitive performance after adjusting for baseline MMSE scores.

Conclusion: AF may accelerate cognitive decline and is associated with greater cortical atrophy in patients with AD. These findings suggest that close cognitive monitoring and individualized care strategies may be beneficial in this population, though further research is needed to confirm this.

Keywords: Alzheimer disease, atrial fibrillation, cognitive decline, cortical atrophy, neuroimaging

INTRODUCTION

Atrial fibrillation (AF) is one of the most common cardiac arrhythmias, affecting millions of individuals worldwide. While AF is a well-established risk factor for stroke and other cardiovascular diseases, recent studies have also highlighted its strong association with cognitive decline and dementia.¹⁻³ The impact of AF on cognitive function extends beyond thromboembolic events and may involve additional mechanisms such as cerebral hypoperfusion, neuroinflammation, and microvascular damage.⁴⁻⁶

Meta-analyses have shown that individuals with AF have a 34% higher risk of developing dementia compared to those without AF.⁷ This risk increases threefold in patients who have experienced a stroke; however, AF has also been linked to cognitive decline in individuals without a history of stroke.⁸ Neuroimaging studies have demonstrated that patients with AF more frequently exhibit structural brain abnormalities, including silent cerebral infarcts, brain atrophy, and cortical thinning. Furthermore, AF has been implicated in Alzheimer's disease-specific pathological processes, such as increased beta-amyloid accumulation.^{1,4} Some hypotheses also suggest a

bidirectional relationship, where amyloid accumulation in the heart, as seen in Alzheimer's disease (AD), may contribute to the development of AF.⁴

Despite the detrimental effects of AF on cognitive function, certain therapeutic interventions may modify this trajectory. Rhythm control strategies, including catheter ablation, have been associated with improved cerebral blood flow and may help slow cognitive decline in patients with AF.⁹ Additionally, oral anticoagulant therapy-particularly direct oral anticoagulants-has emerged as an effective approach to reducing the risk of cognitive impairment in individuals with AF.¹⁰

This exploratory study aims to evaluate whether AF is associated with greater cognitive decline and more advanced cortical atrophy in patients with AD. Cognitive change was assessed prospectively over a six-month follow-up period using the mini-mental state examination (MMSE), while cortical atrophy was evaluated cross-sectionally at baseline using validated visual rating scales. By comparing cognitive and neuroimaging findings between AD patients with and without

AF, the study seeks to clarify the potential contribution of AF to neurodegenerative processes.

METHODS

Participants

The study was approved by the Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital Ethics Committee (Date: 09.04.2025, Decision No: 2025/107). All procedures were conducted in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments. Informed consent was obtained from all participants or their legal guardians.

This study included 64 patients diagnosed with AD who were followed at the Dementia Outpatient Clinic of Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital. The diagnosis of AD was made according to the National Institute on Aging and the Alzheimer's Association (NIA-AA) criteria.¹¹

A total of 800 patients with AD were screened for the presence of AF, and 33 patients with a confirmed diagnosis of AF were included in the study. A control group of 31 AD patients without AF was formed, matched to the AF group in terms of age, sex, and years of education. Only participants with complete data on these demographic variables were included.

Patients were enrolled regardless of the stage of dementia. Written informed consent was obtained from all participants or their legal guardians.

Only patients who had undergone cognitive assessments using the Mini-Mental State Examination (MMSE), Geriatric Depression Scale (GDS), and Geriatric Anxiety Scale (GAS) at the time of diagnosis were included. To ensure valid cognitive and neuroradiological comparisons, only individuals who had undergone magnetic resonance imaging (MRI) were included.

Exclusion criteria were; the presence of neurodegenerative disorders other than AD, history of stroke or large-vessel disease, severe neurological or psychiatric conditions, inability to complete assessments due to significant visual or auditory impairments, history of alcohol or substance abuse, and any disease significantly affecting physical functioning within the last six months.

Demographic data and disease duration were recorded for all patients. The study population was divided into two groups based on the presence or absence of AF. Cognitive status was assessed using the MMSE, while depression and anxiety levels were evaluated using the GDS and GAS. MRI findings were recorded and evaluated using standardized atrophy rating scales.

Data Collection Tools

All assessment tools used in this study have validated Turkish versions.

Mini-Mental State Examination (MMSE): Developed by Folstein et al.¹² (1975) and validated in Turkish,¹³ the MMSE assesses five domains: orientation, registration, attention and calculation, language, and recall. The maximum score is 30.

Scores are interpreted as follows: 0-9 indicates severe, 10-19 moderate, 20-23 mild cognitive impairment, and 24-30 is considered normal.

Geriatric Depression Scale (GDS): The 30-item GDS is a self-report tool used to assess depressive symptoms in older adults.¹⁴ Items are scored 0 or 1, with higher scores indicating more severe depression. Score interpretation: 0-4=normal, 5-8=mild, 9-11=moderate, ≥12=severe depression.

Geriatric Anxiety Inventory (GAI): The GAI evaluates anxiety symptoms in the elderly across somatic, cognitive, and emotional domains.¹⁵ The Turkish version includes 28 items, of which the first 23 are scored on a 4-point Likert Scale (0-3), yielding a total score range of 0-69. Higher scores indicate greater anxiety. Items 24-28 are clinician-rated and not included in the total score.

Radiological Imaging Findings

All imaging was performed using 1.5 Tesla MRI units (GE Signa Explorer; GE Healthcare, Milwaukee, WI, USA). Non-contrast T1-weighted, T2-weighted, and FLAIR sequences were obtained. T1-weighted images (TR/TE: 543/24 ms) were acquired in axial and sagittal planes; T2-weighted images (TR/TE: 5724/102 ms) were acquired in axial and coronal planes. FLAIR images (TR/TE: 8000/86 ms) were obtained in the axial plane.

Koedam, medial temporal atrophy (MTA), and Fazekas scores were evaluated by a neurologist with over five years of clinical experience, who was blinded to patients' demographic and clinical data. Cortical atrophy was evaluated cross-sectionally at baseline, with no follow-up imaging performed; therefore, no conclusions regarding longitudinal progression can be drawn.

Cognitive decline was evaluated separately and prospectively over a six-month observational period using MMSE scores.

Fazekas Grading System: Vascular lesions were evaluated and recorded based on lesion burden.¹⁶ Fazekas scores were used to assess the burden of white matter hyperintensities as a marker of small vessel disease.

• Fazekas grading system:

- **Fazekas 0:** No lesions or a single punctate lesion (white matter hyperintensity).
- **Fazekas 1:** Multiple punctate lesions.
- **Fazekas 2:** Beginning to coalesce lesions (bridging).
- **Fazekas 3:** Large, confluent lesions.

Medial temporal atrophy (MTA) score: The qualitative assessment of the MTA score was based on the degree of atrophy in the choroid fissure, temporal horn, and hippocampal volume in the coronal plane, as described by Frisoni et al.^{17,18}

• Medial temporal atrophy (MTA) scoring system:

- **MTA 0:** Normal choroid fissure, temporal horn, and hippocampal volume.
- **MTA 1:** Mildly enlarged choroid fissure.

- **MTA 2:** Moderately enlarged choroid fissure, with mild enlargement of the temporal horn and a mild reduction in hippocampal volume.
- **MTA 3:** Significantly enlarged choroid fissure, with moderate enlargement of the temporal horn and a moderate reduction in hippocampal volume.
- **MTA 4:** Severely enlarged choroid fissure, significant enlargement of the temporal horn, and a marked reduction in hippocampal volume.

Koedam score: The qualitative assessment of the Koedam score was based on the degree of parietal atrophy in the sagittal plane, as outlined by Koedam et al.,¹⁹ and further supported by subsequent studies.^{20,21}

• **Koedam (posterior/parietal) atrophy score:**

- **Koedam 0:** No sulcal widening or precuneus atrophy.
- **Koedam 1:** Mild sulcal widening and mild precuneus atrophy.
- **Koedam 2:** Marked sulcal widening and moderate parietal cortical atrophy.
- **Koedam 3:** Severe parietal atrophy with a “knife-edge” sulcal appearance.

Study Design

This study employed a combined design approach: cognitive function was assessed longitudinally over a six-month follow-up period using the MMSE, while structural brain imaging data were obtained and analyzed cross-sectionally at baseline. The primary aim was to explore whether AF is associated with greater cognitive decline and/or more advanced cortical atrophy in patients with AD.

Statistical Analysis

Data were analyzed using SPSS version 27.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were first calculated for participants' demographic characteristics using frequency analyses. The normality of continuous variables was assessed using the Shapiro-Wilk test, along with evaluations of skewness and kurtosis (acceptable range: -3 to +3).

For group comparisons, the independent samples t-test was used for normally distributed variables, while the Mann-Whitney U test was applied to non-normally distributed data. Categorical variables were analyzed using the Pearson chi-square test or Fisher's exact test, as appropriate. Descriptive statistics for numerical data included the mean, standard deviation, median, minimum, and maximum values.

Spearman's correlation coefficient was calculated for bivariate correlation analyses. For multivariate analysis, linear regression models were constructed to evaluate the independent effect of atrial fibrillation on cognitive function. A p-value of <0.05 was considered statistically significant for all analyses.

RESULTS

Participant Characteristics

A total of 64 patients with Alzheimer's disease were included in the study. Of these, 33 patients were diagnosed with atrial fibrillation [AF (+)], while 31 patients without AF comprised the control group [AF (-)].

The mean age was 77.21 ± 6.78 years in the AF (+) group and 74.39 ± 9.25 years in the AF (-) group; the difference was not statistically significant ($p > 0.05$). The proportion of female participants was 54.5% ($n=18$) in the AF (+) group and 45.2% ($n=14$) in the AF (-) group, with no significant difference in sex distribution ($p > 0.05$). The mean duration of education was 3.88 ± 3.45 years in the AF (+) group and 3.77 ± 3.01 years in the AF (-) group, also showing no significant difference ($p > 0.05$).

These findings indicate that the two groups were comparable in terms of key demographic characteristics (Table 1).

Table 1. Demographic characteristics of the participants

Characteristic	AF (+) (n=33)	AF (-) (n=31)	Total (n=64)	p-value
Age (years)	77.21 ± 6.78	74.39 ± 9.25	75.83 ± 8.17	>0.05
Sex (female)	18 (54.5%)	14 (45.2%)	32 (50.0%)	>0.05*
Education (years)	3.88 ± 3.45	3.77 ± 3.01	3.83 ± 3.25	>0.05

AF: Atrial fibrillation

At baseline, vascular risk factors were similarly distributed between AF (+) and AF (-) patients. Diabetes mellitus was present in 24.2% of AF (+) patients and 22.6% of AF (-) patients. Hypertension was observed in 66.7% and 58.1%, hyperlipidemia in 18.2% and 25.8%, and current smoking in 3.0% and 12.9% of AF (+) and AF (-) patients, respectively. Regarding body-mass index (BMI), in the AF (+) group 12.5% were normal weight, 37.5% overweight, 45.8% obese, and 4.2% severely obese, while in the AF (-) group the proportions were 25.0%, 37.5%, 29.2%, and 8.3%, respectively.

Cognitive Assessment (MMSE)

At baseline, the difference in MMSE scores between the groups was at the threshold of statistical significance (AF [+]: 20.07 ± 5.80 ; AF [-]: 17.73 ± 4.78 ; $p=0.052$). After six months of follow-up, MMSE scores were 19.64 ± 6.51 in the AF (+) group and 17.64 ± 6.07 in the AF (-) group, with no statistically significant difference between the groups ($p=0.28$).

Within-group changes over time: In the AF (+) group, MMSE scores significantly declined over six months (from 20.58 to 18.55; $p=0.0004$), while no significant change was observed in the AF (-) group (from 18.94 to 18.55; $p=0.5215$). This comparison is illustrated in Figure 1.

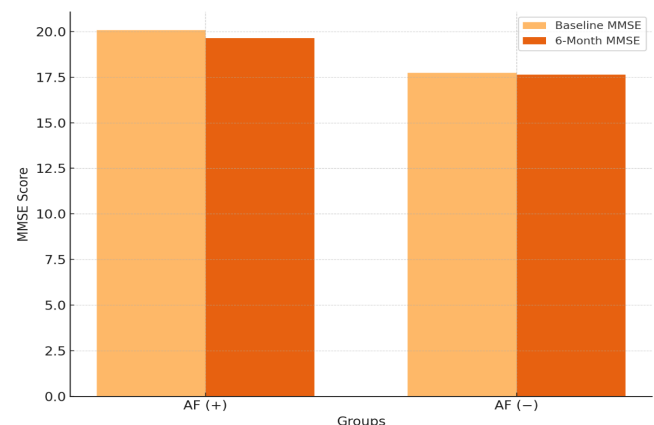


Figure 1. Comparison of MMSE scores at baseline and after 6 months in patients with atrial fibrillation [AF (+)] and without atrial fibrillation [AF (-)]

MMSE: Mini-Mental State Examination, AF: Atrial fibrillation

Radiological Findings

There were no statistically significant differences between the AF (+) and AF (-) groups in terms of Fazekas and Medial Temporal Atrophy (MTA) scores ($p=0.935$ and $p=0.898$, respectively). However, the Koedam score was significantly higher in the AF (+) group (1.85 ± 0.67) than in the AF (-) group (1.55 ± 0.57 ; $p=0.027$), suggesting a possible increase in cortical atrophy in patients with atrial fibrillation (Table 2 and Figure 2).

Table 2. Comparison of radiological atrophy scores between AF (+) and AF (-) groups. A significantly higher Koedam score was observed in the AF (+) group, indicating increased cortical atrophy

MRI Scale	AF (+)	AF (-)	p-value
Fazekas	1.85±1.23	1.74±1.12	0.935
MTA	1.85±1.23	1.74±1.12	0.898
Koedam	1.85±0.67	1.55±0.57	0.027

AF: Atrial fibrillation, MRI: Magnetic resonance imaging, MTA: Medial temporal atrophy

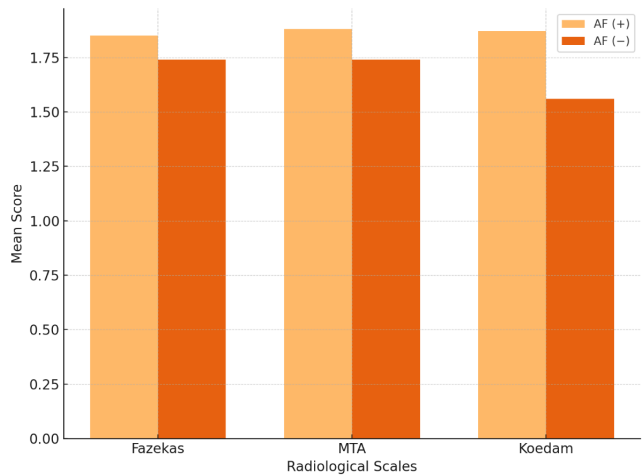


Figure 2. Comparison of fazekas, MTA, and Koedam scores between AF (+) and AF (-) groups in patients with Alzheimer's disease

MTA: Medial temporal atrophy, AF: Atrial fibrillation

Correlation Analyses

No significant correlation was found between MMSE scores and the presence of AF ($p>0.05$).

A weak but statistically significant positive correlation was observed between Koedam scores and the presence of AF.

Both MTA and Fazekas scores demonstrated significant negative correlations with MMSE scores:

- MTA vs. MMSE: $r=-0.500$, $p<0.001$
- Fazekas vs. MMSE: $r=-0.380$, $p=0.002$

Regression Analysis

Multiple linear regression models were constructed to assess the independent effect of AF on cognitive function. After adjusting for baseline MMSE, AF demonstrated a borderline negative association with 6-month cognitive performance, which did not reach conventional levels of statistical significance. Educational level emerged as a positive and statistically significant predictor in all three models ($p<0.05$ in models 1 and 2) (Table 3 and Figure 3).

Table 3. When controlling for baseline MMSE, AF showed a borderline significant negative effect on 6-month cognitive performance. Educational level emerged as a positive and statistically significant predictor in all three models ($p<0.05$ in models 1 and 2)

Model	AF (β)	p-value	Significance
Baseline MMSE	+2.32	0.22	Not significant
6-month MMSE	+0.77	0.68	Not significant
6-month MMSE (adjusted for baseline MMSE)	-1.36	0.108	Borderline significance

AF: Atrial fibrillation; MMSE: Mini-Mental State Examination

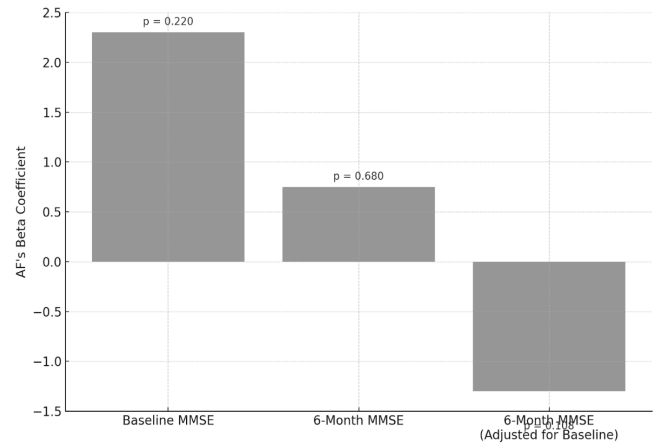


Figure 3. Beta coefficients and corresponding p-values from three regression models showing the independent effect of atrial fibrillation on cognitive performance as measured by the MMSE

MMSE: Mini-Mental State Examination

Adjusted Analysis Including Vascular Risk Factors

In the fully adjusted model including vascular risk factors, the previously significant association between AF and parietal atrophy (Koedam score) was no longer observed. In contrast, the short-term decline in MMSE scores remained borderline significant ($\beta=-1.85$; $p=0.086$), indicating a trend that may warrant further investigation regarding the possible role of AF in cognitive decline. These findings highlight the need for longer-term follow-up to determine whether this trend becomes statistically significant over time.

DISCUSSION

In this study, patients with AD and AF exhibited more pronounced cognitive decline than those without AF, particularly reflected in increased cortical atrophy, as indicated by higher Koedam scores. Based on the MMSE, a significant decline in cognitive performance was observed over six months in the AF (+) group, whereas no such change was detected in the AF (-) group. However, this decline was modest, and regression analyses indicated a borderline association after adjusting for baseline MMSE scores. These findings should be interpreted with caution, in line with the exploratory nature of the study.

Although the initially significant association between AF and parietal atrophy disappeared after adjusting for vascular risk factors, a borderline association with short-term MMSE decline persisted. This trend suggests a possible relationship between AF and cognitive deterioration, but causality cannot be inferred from the current data. Given the small sample size

and short follow-up duration, further longitudinal studies are needed to determine whether AF contributes independently to cognitive decline in Alzheimer's disease.

Our findings raise the possibility that, beyond traditional vascular risk factors, AF may play a role in neurodegeneration; however, this hypothesis requires confirmation in larger samples. These results may support closer monitoring of cognitive function in AD patients with AF, but clinical recommendations should await further evidence.

These findings suggest that AF may contribute to neurodegenerative processes in Alzheimer's disease not only through cardiovascular pathways but also via direct cerebral mechanisms, although confirmation through longer follow-up with serial imaging is needed. In the literature, the link between AF and cognitive impairment has been attributed to several factors, including cerebral hypoperfusion, neuroinflammation, microvascular damage, and beta-amyloid accumulation.^{7,22-24} Silent brain infarcts and microemboli are also recognized as major contributors to cognitive decline in individuals with AF.

In our study, the significant association between AF and cortical atrophy (as measured by the Koedam score) supports the involvement of these neurodegenerative mechanisms. However, this association was no longer observed after adjusting for vascular risk factors, suggesting it may be confounded by comorbidities. Previous studies have reported a higher frequency of posterior cortical atrophy in patients with AF, which correlates with greater cognitive decline.²⁵ Additionally, large-scale cohort studies have demonstrated strong associations between AF, silent brain lesions, and lower cognitive performance.²⁶

Our results also revealed a positive and significant relationship between educational level and MMSE scores, supporting the theory of "cognitive reserve." Higher educational attainment may confer resilience against cognitive decline in the presence of risk factors such as AF.^{27,28}

An important aspect in the literature regarding AF and cognitive outcomes is the role of treatment strategies. In a meta-analysis conducted by Chen et al.,²⁹ catheter ablation was shown to reduce the risk of dementia by 40% and significantly improve MMSE scores in patients with AF. While these findings are promising, their generalizability to patients with comorbid AD remains unclear. Similarly, studies have demonstrated that oral anticoagulant (OAC) therapy—particularly with direct oral anticoagulants (DOACs)—significantly reduces the risk of cognitive impairment.^{30,31} In our study, only one patient received ablation therapy, and data regarding OAC or DOAC use did not reach statistical significance due to the small sample size. Future studies should more robustly investigate the cognitive effects of specific AF treatments in AD populations.

The strengths of this study include the selection of both AF (+) and AF (-) groups from a population with confirmed AD diagnoses, ensuring diagnostic homogeneity and enabling robust group comparisons. The combined analysis of cognitive and radiological data, along with the use of multivariate regression models, allowed for a more controlled evaluation of AF's possible impact on cognitive function.

Nevertheless, the study is limited by its modest sample size, short follow-up duration, and lack of comprehensive data on AF subtype or anticoagulant use. These constraints limit the generalizability of our findings and underscore the need for larger, longer-term prospective studies.

Limitations

This study has several limitations. First, the relatively small sample size and the short six-month follow-up limit the generalizability and longitudinal interpretation of the findings. Second, key clinical parameters such as AF subtype (paroxysmal, persistent, or permanent), anticoagulant use, and AF duration were not comprehensively analyzed due to limited data availability. These factors may have influenced the results and should be considered in future studies.

In addition, the use of a single cognitive screening tool (MMSE) limits the ability to detect more subtle or domain-specific cognitive changes. Future research should employ comprehensive neuropsychological assessments that evaluate multiple cognitive domains, including attention, memory, and executive function.

Long-term, multicenter prospective studies with larger cohorts are needed to better clarify the independent effects of AF on cognitive decline in patients with Alzheimer's disease. Incorporating functional imaging techniques—such as fMRI and PET—may also enhance understanding of the pathophysiological links between AF and neurodegeneration. Finally, future studies should explore the differential cognitive effects of specific oral anticoagulants (e.g., DOACs vs. vitamin K antagonists) to inform more personalized treatment approaches.

CONCLUSION

This study investigated the potential impact of AF on cognitive function and structural brain changes in patients with Alzheimer's disease. The findings indicated that cognitive decline was more pronounced in Alzheimer's patients with AF, particularly in relation to cortical atrophy, as evidenced by significantly higher Koedam scores. A statistically significant reduction in MMSE scores was observed over a six-month period in the AF (+) group, whereas no such decline occurred in the AF (-) group. Additionally, regression analyses revealed a borderline significant negative effect of AF on cognitive performance after adjusting for baseline MMSE scores. These results suggest that AF may be associated with more pronounced cortical atrophy and cognitive decline in Alzheimer's disease, although confirmation through larger and longer-term studies is needed. Given the exploratory nature of the study and the limited sample, these findings should be considered hypothesis-generating rather than conclusive. Monitoring cognitive function in patients with both AD and AF may be important, but treatment decisions should be based on a comprehensive clinical evaluation.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital Ethics Committee (Date: 09.04.2025, Decision No: 2025/107).

Informed Consent

Written informed consent was obtained from all participants or their legal guardians.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Data, Material and/or Code Availability

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

Acknowledgments

The authors would like to thank the patients and their families for their participation in this study. We also acknowledge the support of the clinical staff of the Dementia Outpatient Clinic at Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital. The authors declare that they received no financial support for this study and that there are no conflicts of interest.

REFERENCES

- Huang LJ, Chen JS, Song YZ, Chang P. Atrial fibrillation and cognitive impairment: mechanisms, influencing factors, and prospects. *Front Cardiovasc Med*. 2025;12:1527802. doi:10.3389/fcvm.2025.1527802
- Militaru M, Lighezan DF, Tudoran C, Militaru AG. Connections between cognitive impairment and atrial fibrillation in patients with diabetes mellitus type 2. *Biomedicines*. 2024;12(3):672. doi:10.3390/biomedicines12030672
- Zenger B, Rizzi S, Steinberg BA, Ranjan R, Bunch TJ. This is your brain, and this is your brain on atrial fibrillation: the roles of cardiac malperfusion events and vascular dysfunction in cognitive impairment. *Arrhythm Electrophysiol Rev*. 2023;12:e01. doi:10.15420/aer.2022.29
- Morales-Bacas E, Duque-Holguera M, Portilla-Cuenca JC, Casado-Naranjo I. Atrial fibrillation and cognitive impairment: a narrative review. *Rev Neurol*. 2022;75(10):311-318. doi:10.33588/rn.7510.2022252
- Kelley RE, Kelley BP. Heart-brain relationship in stroke. *Biomedicines*. 2021;9(12):1835. doi:10.3390/biomedicines9121835
- Hu Z, Ding L, Yao Y. Atrial fibrillation: mechanism and clinical management. *Chin Med J (Engl)*. 2023;136(22):2668-2676. doi:10.1097/CM9.00000000000002906
- Papanastasiou CA, Theochari CA, Zareifopoulos N, et al. Atrial fibrillation is associated with cognitive impairment, all-cause dementia, vascular dementia, and Alzheimer's disease: a systematic review and meta-analysis. *J Gen Intern Med*. 2021;36(10):3122-3135. doi:10.1007/s11606-021-06954-8
- Filler J, Georgakis MK, Dichgans M. Risk factors for cognitive impairment and dementia after stroke: a systematic review and meta-analysis. *Lancet Healthy Longev*. 2024;5(1):e31-e44. doi:10.1016/S2666-7568(23)00217-9
- Yoshihisa A, Kono S, Kaneshiro T, et al. Impaired brain activity in patients with persistent atrial fibrillation assessed by near-infrared spectroscopy and its changes after catheter ablation. *Sci Rep*. 2022;12(1):7866. doi:10.1038/s41598-022-12097-5
- Grymonprez M, Petrovic M, De Backer TL, Ikram MA, Steurbaut S, Lahousse L. Comparing the risk of dementia in subjects with atrial fibrillation using non-vitamin K antagonist oral anticoagulants versus vitamin K antagonists: a Belgian nationwide cohort study. *Age Ageing*. 2023;52(3):afad038. doi:10.1093/ageing/afad038
- McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263-269. doi:10.1016/j.jalz.2011.03.005
- Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *J Psychiatr Res*. 1975;12(3):189-198. doi:10.1016/0022-3956(75)90026-6
- Güngen C, Ertan T, Eker E, Yaşar R, Engin F. Reliability and validity of the standardized mini mental state examination in the diagnosis of mild dementia in Turkish population. *Türk Psikiyatri Derg Turk J Psychiatry*. 2002;13(4):273-281.
- Yesavage JA, Brink TL, Rose TL, et al. Development and validation of a Geriatric Depression Screening Scale: a preliminary report. *J Psychiatr Res*. 1982;17(1):37-49. doi:10.1016/0022-3956(82)90033-4
- Pifer MA, Segal DL. Geriatric Anxiety Scale: development and preliminary validation of a long-term care anxiety assessment measure. *Clin Gerontol*. 2020;43(3):295-307. doi:10.1080/07317115.2020.1725793
- Fazekas F, Chawluk JB, Alavi A, Hurtig HI, Zimmerman RA. MR signal abnormalities at 1.5 T in Alzheimer's dementia and normal aging. *AJR Am J Roentgenol*. 1987;149(2):351-356. doi:10.2214/ajr.149.2.351
- Frisoni GB, Scheltens Ph, Galluzzi S, et al. Neuroimaging tools to rate regional atrophy, subcortical cerebrovascular disease, and regional cerebral blood flow and metabolism: consensus paper of the EADC. *J Neurol Neurosurg Psychiatry*. 2003;74(10):1371-1381. doi:10.1136/jnnp.74.10.1371
- Wahlund LO, Westman E, van Westen D, et al. Imaging biomarkers of dementia: recommended visual rating scales with teaching cases. *Insights Imaging*. 2017;8(1):79-90. doi:10.1007/s13244-016-0521-6
- Koedam EL, Lehmann M, van der Flier WM, et al. Visual assessment of posterior atrophy development of a MRI Rating Scale. *Eur Radiol*. 2011;21(12):2618-2625. doi:10.1007/s00330-011-2205-4
- Furtner J, Prayer D. Neuroimaging in dementia. *Bildgebung in der Demenzdiagnostik. Wien Med Wochenschr*. 2021;171(11-12):274-281. doi:10.1007/s10354-021-00825-x
- Raposo Rodríguez L, Tovar Salazar DJ, Fernández García N, Pastor Hernández L, Fernández Guinea Ó. Magnetic resonance imaging in dementia. *Resonancia magnética en las demencias. Radiología (Engl Ed)*. 2018;60(6):476-484. doi:10.1016/j.rx.2018.04.003
- Kokkinidis DG, Zareifopoulos N, Theochari CA, et al. Association between atrial fibrillation and cognitive impairment in individuals with prior stroke: a meta-analysis and meta-regression analysis. *Stroke*. 2020;51(6):1662-1666. doi:10.1161/STROKEAHA.119.027815
- Chang YT, Chen YL, Kang HY. Revealing the influences of sex hormones and sex differences in atrial fibrillation and vascular cognitive impairment. *Int J Mol Sci*. 2021;22(16):8776. doi:10.3390/ijms22168776
- Fekete M, Liotta EM, Molnar T, Füllöp GA, Lehoczki A. The role of atrial fibrillation in vascular cognitive impairment and dementia: epidemiology, pathophysiology, and preventive strategies. *Geroscience*. 2025;47(1):287-300. doi:10.1007/s11357-024-01290-1
- Schwenneken H, Browndyke JN, Wright MC, et al. A pilot study of longitudinal changes in neurocognition, white matter hyperintensities, and cortical thickness in atrial fibrillation patients following catheter ablation vs medical management. *Heart Rhythm O2*. 2024;5(2):122-130. doi:10.1016/j.hroo.2024.01.002
- D'Anna L, Filippidis FT, Harvey K, et al. Extent of white matter lesion is associated with early hemorrhagic transformation in acute ischemic stroke related to atrial fibrillation. *Brain Behav*. 2021;11(8):e2250. doi:10.1002/brb3.2250
- Mailhot T, McManus DD, Waring ME, et al. Frailty, Cognitive impairment, and anticoagulation among older adults with nonvalvular atrial fibrillation. *J Am Geriatr Soc*. 2020;68(12):2778-2786. doi:10.1111/jgs.16756
- Madhavan M, Holmes DN, Piccini JP, et al. Association of frailty and cognitive impairment with benefits of oral anticoagulation in patients with atrial fibrillation. *Am Heart J*. 2019;211:77-89. doi:10.1016/j.ahj.2019.01.005

29. Chen P, Chen Z, Pan D, et al. Catheter ablation and cognitive function in atrial fibrillation: a systematic review and meta-analysis. *Front Neurol.* 2022;13:934512. doi:10.3389/fneur.2022.934512.
30. Agarwal A, Mostafa MA, Ahmad MI, Soliman EZ. Exploring the link between anticoagulation, cognitive impairment and dementia in atrial fibrillation: a systematic review. *J Clin Med.* 2024;13(8):2418. doi:10.3390/jcm13082418
31. Chung SC, Rossor M, Torralbo A, et al. Cognitive impairment and dementia in atrial fibrillation: a population study of 4.3 million individuals. *JACC Adv.* 2023;2(9):100655. doi:10.1016/j.jacadv.2023.100655

Evaluation of clinical and laboratory findings in cases of infantile hypertrophic pyloric stenosis

 Fatih Kurt¹,  Murat Kaya²

¹Department of Pediatrics, Faculty of Medicine, Düzce University, Düzce, Türkiye

²Department of Pediatric Surgery, Faculty of Medicine, Düzce University, Düzce, Türkiye

Cite this article as: Kurt F, Kaya M. Evaluation of clinical and laboratory findings in cases of infantile hypertrophic pyloric stenosis. *J Health Sci Med.* 2025;8(6):1030-1033.

Received: 28.08.2025

Accepted: 24.09.2025

Published: 25.10.2025

ABSTRACT

Aims: Infantile hypertrophic pyloric stenosis (IHPS) is the most common gastrointestinal disorder in early infancy that requires surgical intervention. Although its etiology is not fully understood, several risk factors have been identified, including male sex, bottle feeding, and being a first-born child. The aim of this study was to evaluate the clinical and laboratory findings of patients with IHPS and to contribute to the existing literature.

Methods: This study retrospectively analyzed 19 cases of IHPS diagnosed from 2012 to 2024 at the Department of Pediatrics, Düzce University Research and Training Hospital. Demographic characteristics, presenting complaints, laboratory results, and birth histories of the patients were analyzed. Descriptive statistics and correlation analyses were performed using SPSS version 25.

Results: Of the patients, 78.9% were male, with a median age at diagnosis of 42 days. Bottle feeding was observed in 47.4% of cases, and 47.4% were first-born children. Vomiting was present in all patients, while 68.4% had complaints of poor weight gain. The rate of hypochloremia was 42.1%. A moderate negative correlation was found between maternal age and the duration of symptoms.

Conclusion: Early recognition is crucial in the diagnosis and management of IHPS. Training healthcare professionals on this condition may improve early diagnosis and treatment outcomes. Our study provides valuable insights into the risk factors and clinical characteristics of IHPS.

Keywords: Infantile hypertrophic pyloric stenosis, metabolic alkalosis, olive sign, nonbilious vomiting

INTRODUCTION

Infantile hypertrophic pyloric stenosis (IHPS) is characterized by pathological thickening of the pyloric sphincter muscle, leading to gastric outlet obstruction. It was first identified by Blair in 1717 during a postmortem study and later described in detail by Hirschsprung in 1888, who proposed it as a congenital involution of the fetal pylorus. Symptoms typically manifest between the second and eighth weeks of life, presenting as non-bilious, projectile vomiting. Dehydration, hypochloremic-hypokalemic metabolic alkalosis and growth retardation are the accompanying signs. Increased gastric peristaltic movements and an olive-shaped mass in the epigastric region can be detected during physical examination. The diagnosis is confirmed through ultrasonography.^{1,2} The most widely accepted diagnostic criteria for pyloric stenosis on ultrasonographic evaluation are a pyloric muscle thickness of >4 mm and a pyloric length of >16 mm. The combined use of these criteria has been reported in the literature to provide high diagnostic accuracy, with a sensitivity of 94% and a specificity of 91.7%. In contrast, the sensitivity of detecting the

“olive” sign by palpation has been reported to vary between 10% and 93%.³

The incidence of IHPS is approximately 2 per 1,000 live births. Studies conducted in the 1990s reported a rising incidence; however, more recent investigations suggest a declining trend. IHPS is approximately five times more common in males compared to females.⁴ While the diagnosis and treatment of IHPS are well-established, its etiology remains incompletely understood. Identified risk factors include male gender, maternal smoking, bottle feeding, cesarean delivery, African ancestry, and being the first-born child.⁵

IHPS is the most common gastrointestinal disorder requiring surgical intervention in the early stages of life. Historically, the mortality rate reached up to 50%; however, with advancements in surgical techniques and medical management, mortality and morbidity rates have significantly decreased.⁶ The treatment gold standard is laparoscopic pyloromyotomy, a minimally invasive procedure. Studies have identified poor

prognostic factors, including symptom onset before two weeks of age, delayed presentation, prolonged preoperative hospitalization, electrolyte imbalances, severe dehydration, and surgical site infections.⁷

In this study, we aimed to analyze the preoperative clinical and laboratory characteristics of IHPS cases diagnosed at our center over the past 12 years and to identify potential risk factors, thereby contributing valuable insights to the existing body of literature.

METHODS

The study was initiated with the approval of the Düzce University Faculty of Medicine Clinical Researches Ethics Committee (Date: 16.12.2024, Decision No: 2024/257). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

This retrospective study evaluated 19 cases diagnosed with IHPS between 2012 and 2024 at the Pediatrics Department of Düzce University Faculty of Medicine. The cases were analyzed based on demographic and clinical parameters, including age at admission, gender, weight, gestational age, birth weight, mode of delivery, birth order, feeding method (bottle-fed or not), presenting symptoms, presence of polyhydramnios during pregnancy, maternal age, physical examination findings, associated anomalies, laboratory results, and radiological imaging data.

Statistical Analysis

Descriptive data analyses of the data were presented as frequencies and percentages. For normality testing, when the sample size was greater than 50, the bell curve observation was conducted, and when the sample size was less than 50, the Shapiro-Wilk test was used. Parameters with normal distribution were presented as mean±standard deviation (SD), while parameters without normal distribution were presented as median (interquartile range: 25-75). All statistical analyses were performed using SPSS for Windows, version 25 (IBM, Chicago, IL, USA). Pearson correlation analysis was applied to evaluate the correlation between maternal age and symptom duration. The Pearson correlation coefficient (r) was interpreted as follows: values below 0.20 were considered negligible correlation, 0.20-0.50 as weak correlation, 0.50-0.70 as moderate correlation, and values above 0.70 as strong correlation (8). The correlation coefficients (r) and significance values (p) were reported, with results considered statistically significant at p<0.05.

RESULTS

A total of 19 cases were diagnosed with IHPS at our hospital over the past 12 years. The median age of the patients was 42 days, with 15 (78.9%) being male. The mean gestational age at birth was 37.11±2.58 weeks, and the mean birth weight was 2832.11±608.01 grams. Ten (52.6%) of the cases were delivered via normal spontaneous vaginal delivery, and 11 (57.9%) had a history of bottle feeding. When inquired about birth order, 9 (47.4%) were the firstborn, with the median being the second child. Vomiting was present in all of the cases, with 14 (73.7%) showing the olive sign on palpation, and 13 (68.4%) exhibiting

inadequate weight gain. The median onset of symptoms was 27 days (range 15-30), and the average maternal age was 28.78±5.30 years. One case had a family history of IHPS in both the mother and maternal uncle. One patient had a comorbid condition of congenital hypothyroidism. Among the mothers, 63.2% were primary school graduates, 31.5% were secondary school graduates, and only one case had a university degree. Additionally, 26.3% of the mothers were found to be smokers (Table 1).

Table 1. Some sociodemographic data and clinical characteristics of the cases

Variable	n=19 (%)
Age (day) [median (IQR 25-75)]	42 (32-67)
Gender	
Boy	15 (78.9)
Girl	4 (21.1)
Gestational age (weeks) (mean±SD)	37.11±2.58
Prematurity (<37 weeks)	6 (31.5)
Birth weight (grams) (mean±SD)	2832.11±608.01
Mode of delivery	
C/S	9 (47.4)
NSVD	10 (52.6)
Parity [median (IQR 25-75)]	2 (1-3)
Firstborn child	9 (47.4)
Multiple pregnancy	1 (5.3)
Bottle feeding	11 (57.9)
Familial IHPS	1* (5.3)
Mother's education level	
Secondary school graduate	12 (63.2)
High school graduate	6 (31.5)
University graduate	1 (5.3)
Mother smoking	
Yes	5 (26.3)
No	14 (73.7)
Symptoms and signs	
Olive sign	14(73.7)
Vomiting	19 (100)
Inability to gain weight	13 (68.4)
Time to onset of symptoms (days) [median (IQR 25-75)]	27 (15-30)
Additional disease	1 (5.3)
Maternal age (years)	28.78±5.30

* This patient has a family history of IHPS in both the mother and the maternal uncle. IQR: Inter quartile range, SD: Standard deviation, C/S: Caesarean section, NSVD: Normal spontaneous vaginal delivery, IHPS: Infantile hypertrophic pyloric stenosis

The mean hemoglobin level in the cases was 10.65±0.94 g/dl, and the hematocrit level was 32.24±4.46%. Blood gas analysis revealed a mean pH of 7.46±0.05, pCO₂ of 41.60±5.93 mmHg, median HCO₃⁻ of 26.4 (25.9-30.9) mEq/L, base deficit of 2.9 (1.7-8.5), and lactate level of 2.1 (1.5-2.8). The mean plasma sodium concentration was 135.56±3.19 mEq/L, potassium 4.53±0.75 mEq/L, chloride 98.07±6.91 mEq/L, and calcium 10.06±0.43 mg/dl. Furthermore, the mean plasma glucose

level was 87.47±13.24 mg/dl, median urea 10.9 (6.9-20) mg/dl, BUN 5.3 (3-8) mg/dl, and creatinine 0.32 (0.27-0.41) mg/dl (Table 2).

Table 2. Some laboratory parameters of the cases	
Laboratory parameters	Mean±SD
Hemoglobin (g/dl)	10.65±0.94
Hematocrit (%)	32.24±4.46
Blood gas pH level	7.46±0.05
pCO2 level (mmHg)	41.60±5.93
HCO3 level (mEq/L) [median (IQR 25-75)]	26.4 (25.9-30.9)
Base deficit (mmol/L) [median (IQR 25-75)]	2.9 (1.7-8.5)
Blood lactate (mg/dl) [median (IQR 25-75)]	2.1 (1.5-2.8)
Plasma sodium (mEq/L)	135.56±3.19
Plasma potassium (mEq/L)	4.53±0.75
Plasma calcium (mg/dl)	10.06±0.43
Plasma chloride (mEq/L)	98.07±6.91
Plasma glucose (mg/dl)	87.47±13.24
Plasma urea (mg/dl) [median (IQR 25-75)]	10.9 (6.9-20)
BUN (mg/dl) [median (IQR 25-75)]	5.3 (3-8)
Plasma creatinin (mg/dl) [median (IQR 25-75)]	0.32 (0.27-0.41)
IQR: Inter quartile range, SD: Standard deviation	

A moderate negative correlation (r=-0.496) was identified between maternal age and symptom duration, and this relationship was found to be statistically significant (p=0.031) (Table 3).

Table 3. The relationship between maternal age and symptom duration		
		Symptom duration
Maternal age	Pearson r	-0.496*
	p	0.031
	n	19
*The correlation is significant at the level of p=0.031		

DISCUSSION

IHPS is a common clinical condition affecting young infants. Although its etiology is not fully understood, advances in medical and surgical techniques have reduced the mortality rate from 50% to nearly 0%.⁶ One of the most significant risk factors for IHPS is male sex, with a reported male-to-female ratio ranging from 4:1 to 6:1.⁹ Consistent with this, 78.9% of the cases in our study were male.

Ultrasonography is utilized as a reliable method for the diagnosis of pyloric stenosis, demonstrating high sensitivity and specificity. However, diagnostic rates based on physical examination findings have been reported to range from 10% to 93% and have exhibited a notable decline in recent years.³ In our study, the detection rate of physical examination findings was also low. The development and widespread availability of modern imaging techniques such as ultrasound may have contributed to a decline in clinicians' physical examination skills.

In the literature, the average age at diagnosis of IHPS has been reported to range between 36.5 and 43.1 days.⁹ Studies have shown that diagnosis is delayed by more than one week in one-third of cases, and such delays are associated with more severe metabolic disturbances and decreased surgical success rates.¹⁰ In our study, the age of symptom onset and the age at diagnosis of the cases were consistent with the literature. However, in our patient population, cases where diagnosis was delayed up to 87 days were also observed. This delay is likely due to challenges in differentiating IHPS from gastroesophageal reflux and delays in performing diagnostic ultrasonography. Increasing healthcare professionals' awareness of IHPS could significantly improve early diagnosis and treatment outcomes.

IHPS typically manifests between the 3rd and 8th weeks of life. Prolonged vomiting leads to dehydration and inadequate weight gain in affected infants.¹¹ Taylor et al.,¹² in a study of 362 cases, reported that vomiting was present in 99.7% of cases and failure to gain weight in 46%. Bundak et al.¹³ stated that an infant should gain an average of 20-30 grams per day during the first three months of life. In our study, all cases presented with vomiting, and 68.4% of the infants failed to achieve the daily baseline average weight gain of 20 grams. The high rate of inadequate weight gain in our study may be attributed to the limited sample size.

In the study conducted by Chuang et al.⁹ on the European population, a family history was identified in 14.7% of cases; however, they reported that family history was rarely encountered in their cohort. Additionally, it was stated that accompanying gastrointestinal system anomalies were also rare. In our study, while a family history of IHPS was present in the mother and uncle of one case, no accompanying gastrointestinal system anomalies were detected in any of our cases.

Heterogeneous data regarding gestational age and prematurity have been reported in the literature.⁷ In a study conducted in Türkiye by Kaya et al.,¹⁴ the mean gestational age of the cases was reported as 38±2.3 weeks, with 14.7% being premature and 50.8% delivered via C/S. In our study, a high prematurity rate, which we consider specific to our cohort, was observed. This finding may be attributed to the relatively small sample size of our study. On the other hand, the mean gestational age and the C/S rate in our study were similar to those reported by Kaya et al.

In a meta-analysis conducted by Zhu et al.,¹⁵ it was reported that the risk of IHPS is higher in firstborn children and is also slightly increased in multiple pregnancies. In our study, 9 cases (47.4%) were firstborn children, and only 1 case was born from a twin pregnancy.

Bottle feeding has been reported as a major risk factor for IHPS, increasing its incidence by 4.6 times; while breast milk relaxes the pyloric muscle and accelerates gastric emptying, formula feeding delays emptying and contributes to pylorospasm and hypertrophy.^{7,16} In our study, 47.4% of cases had a history of bottle feeding and our study also supports its role as a significant risk factor.

Studies have reported that hypochloremia and alkalosis are independent predictors of IHPS. In a study by Gotley et al.,¹⁷

alkalosis was detected in 61% of cases, while hypochloremia was observed in 29%. In our study, hypochloremia was observed in 42.1% of the cases, and alkalosis was present in 57.8%. The difference in hypochloremia is likely related to the duration of symptoms.

Although Chuang et al.⁹ and Feng et al.¹⁸ reported associations between symptom duration and bicarbonate or electrolyte levels, no such relationship was observed in our study. However, a moderate negative correlation was identified between maternal age and symptom duration, suggesting that older mothers may recognize abnormalities in their children earlier and seek medical attention more promptly. This observation is supported by a study by Broom et al.,¹⁹ which highlights maternal age as one of the two most influential factors affecting maternal sensitivity.

Limitations

This study has several limitations. First, the small sample size restricts the generalizability of the results. Second, as the study was retrospective, there is a possibility of incomplete or insufficient data. Third, long-term outcomes were not evaluated, thus preventing insights into the prognosis of IHPS after surgery. Addressing these limitations will require future larger-scale, prospective studies.

CONCLUSION

Our study supports that male gender, bottle feeding, and being a firstborn child are significant risk factors for IHPS. Late recognition of symptoms delays diagnosis and treatment, while increasing maternal age enhances the likelihood of early presentation. Hypochloremia and alkalosis are prominent laboratory findings in IHPS. Raising awareness among healthcare professionals and achieving early diagnosis will positively influence treatment outcomes.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Düzce University Faculty of Medicine Clinical Researches Ethics Committee (Date: 16.12.2024, Decision No: 2024/257).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Zampieri N, Corato V, Scirè G, Camoglio FS. Hypertrophic pyloric stenosis: 10 years' experience with standard open and laparoscopic approach. *Pediatr Gastroenterol Hepatol Nutr*. 2021;24(3):265-272. doi:10.5223/pghn.2021.24.3.265
2. Sakharkar S, Tiwari D, Yemul R, Kabra N. The enigma of infantile hypertrophic pyloric stenosis in preterm neonate. *Clin Pediatr (Phila)*. 2025;99228251342391. doi:10.1177/00099228251342391
3. van den Bunder FA, Derikx JP, Kiblawi R, van Rijn RR, Dingemans J. Diagnostic accuracy of palpation and ultrasonography for diagnosing infantile hypertrophic pyloric stenosis: a systematic review and meta-analysis. *Br J Radiol*. 2022;95(1139):20211251. doi:10.1259/bjr.20211251
4. Oetzmann von Sochaczewski C, Muensterer OJ. The incidence of infantile hypertrophic pyloric stenosis nearly halved from 2005 to 2017: analysis of German administrative data. *Pediatr Surg Int*. 2021;37(5):579-585. doi:10.1007/s00383-020-04810-0
5. Obaid YY, Toubasi AA, Albustanji FH, Al-Qawasmeh AR. Perinatal risk factors for infantile hypertrophic pyloric stenosis: a systematic review and meta-analysis. *J Pediatr Surg*. 2023;58(3):458-466. doi:10.1016/j.jpedsurg.2022.08.016
6. Hernanz-Schulman M. Infantile hypertrophic pyloric stenosis. *Radiology*. 2003;227(2):319-331. doi:10.1148/radiol.2272011329
7. Muse AL, Hussein BO, Adem BM, Osman MO, Abdulahi ZB, Ibrahim MA. Treatment outcome and associated factors of infantile hypertrophic pyloric stenosis at eastern Ethiopia public hospitals. *BMC Surg*. 2024;24(1):262. doi:10.1186/s12893-024-02567-0
8. Kozak M. What is strong correlation? *Teaching Statistics*. 2009;31(3):85-86. doi:10.1111/j.1467-9639.2009.00387.x
9. Chuang YH, Chao HC, Yeh HY, Lai MW, Chen CC. Factors associated with pyloric hypertrophy severity and post-operative feeding and nutritional recovery in infantile hypertrophic pyloric stenosis. *Biomed J*. 2022;45(6):948-956. doi:10.1016/j.bj.2021.12.011
10. Acker SN, Garcia AJ, Ross JT, Somme S. Current trends in the diagnosis and treatment of pyloric stenosis. *Pediatr Surg Int*. 2015;31(4):363-366. doi:10.1007/s00383-015-3682-3
11. Fawcner-Corbett D, McHoney M. Infantile hypertrophic pyloric stenosis. *Paediatr Child Health*. 2024;34(4):126-129. doi:10.1016/j.paed.2024.01.004
12. Taylor ND, Cass DT, Holland AJ. Infantile hypertrophic pyloric stenosis: has anything changed? *J Paediatr Child Health*. 2013;49(1):33-37. doi:10.1111/jpc.12027
13. Bundak R, Neyzi O, Günöz H, Darendeliler F. *Pediatric Cilt 1. 3. Baskı. Nobel Tıp Kitabevleri*; 2002. Büyüme-gelişme ve bozuklukları; 79-99.
14. Kaya B, Akduman H, Dilli D, et al. Pyloric Index, a new parameter: predicting perioperative prognosis in neonates with hypertrophic pyloric stenosis. *Scand J Gastroenterol*. 2024;59(12):1272-1276. doi:10.1080/00365521.2024.2425982
15. Zhu J, Zhu T, Lin Z, Qu Y, Mu D. Perinatal risk factors for infantile hypertrophic pyloric stenosis: a meta-analysis. *J Pediatr Surg*. 2017;52(9):1389-1397. doi:10.1016/j.jpedsurg.2017.02.017
16. Krogh C, Biggar RJ, Fischer TK, et al. Bottle-feeding and the risk of pyloric stenosis. *Pediatrics*. 2012;130(4):e943-e949. doi:10.1542/peds.2011-2785
17. Gotley LM, Blanch A, Kimble R, Frawley K, Acworth JP. Pyloric stenosis: a retrospective study of an Australian population. *Emerg Med Australas*. 2009;21(5):407-13. doi:10.1111/j.1742-6723.2009.01218.x
18. Feng Z, Nie Y, Zhang Y, et al. The clinical features of infantile hypertrophic pyloric stenosis in Chinese Han population: analysis from 1998 to 2010. *PLoS One*. 2014;9(2):e88925. doi:10.1371/journal.pone.0088925
19. Broom BL. Impact of marital quality and psychological well-being on parental sensitivity. *Nurs Res*. 1994;43(3):138-143. doi:10.1371/journal.pone.0088925

Mothers' knowledge and attitudes regarding circumcision and the influence of parental demographics: a retrospective study

 Damla Erden,  Vefik Arica

Department of Pediatrics, İstanbul Okmeydanı Training and Research Hospital, İstanbul, Türkiye

Cite this article as: Erden D, Arica V. Mothers' knowledge and attitudes regarding circumcision and the influence of parental demographics: a retrospective study. *J Health Sci Med.* 2025;8(6):1034-1040.

Received: 05.07.2025

Accepted: 25.09.2025

Published: 25.10.2025

ABSTRACT

Aims: Circumcision is a surgical procedure performed on nearly all males in our society. This study aimed to evaluate mothers' knowledge and attitudes regarding circumcision and to raise awareness of common misconceptions that may lead to adverse outcomes.

Methods: This retrospective, cross-sectional study was conducted between February 17 and April 17, 2014, among mothers whose children had previously undergone circumcision and who presented to the pediatric outpatient clinic for any reason. A total of 540 mothers were included. A structured questionnaire assessing socioeconomic status, educational background, and knowledge about circumcision was administered.

Results: The study included 540 mothers. The mean age at circumcision was 3.7 ± 2.6 years. Circumcision was performed during the neonatal period in 13.8% (n=75), at 1-2 years in 24.9% (n=135), at 3-6 years in 44.3% (n=240), and after 6 years in 17% (n=92) of cases. Mothers with lower socioeconomic and educational levels had significantly less knowledge regarding circumcision.

Conclusion: The majority of circumcisions were performed between the ages of 3 and 6, a developmentally vulnerable period. The findings suggest that families were generally unaware of the potential psychological effects and medical risks associated with circumcision at this age. Mothers with low income and education levels should be included in educational programs to improve awareness about the benefits and possible complications of circumcision.

Keywords: Circumcision, level of knowledge, complications, maternal attitude

INTRODUCTION

Circumcision is the surgical removal of the prepuce, also known as the foreskin, which covers the glans penis, in a specific shape and length, thereby exposing the tip of the penis.^{1,2} It is a tradition that has been practiced for thousands of years and is considered the most commonly performed surgical procedure worldwide. The first circumcision is believed to have taken place approximately 15,000 years ago. Routine neonatal circumcision began in the 19th century as a preventive measure against diseases and was rapidly adopted, especially in English-speaking countries. Today, neonatal circumcision is a widely performed operation in the United States.^{2,3} Globally, approximately 25% of males are circumcised, while in Türkiye, nearly all adult males are circumcised due to traditional, cultural, and religious reasons.^{3,4}

In our society, circumcision is primarily performed for religious reasons, but it is also perceived as a rite of passage into manhood.⁴ Recently, some have argued that circumcision constitutes a non-consensual intervention on the child's body and disrupts bodily integrity.⁵

There are studies reporting that neonatal circumcision, when performed using appropriate techniques and by specialists, heals more easily and results in fewer complications compared to circumcision in infancy.^{6,7}

Although circumcision is often considered a simple surgical procedure, when performed by untrained individuals, it may result in serious and undesirable complications. These include early complications such as surgical site infections, insufficient skin removal, excessive excision of the prepuce, glans injury, iatrogenic hypospadias, glanular necrosis, amputation, and bleeding that can be fatal in patients with coagulation disorders. Additionally, late complications such as penile adhesions, epidermal inclusion cysts, urethral fistula, buried penis, meatitis and meatal stenosis may also occur.^{8,9}

Circumcision is a commonly practiced surgical procedure in our country. Most male children are circumcised between the ages of 3 and 7, while only a small portion of hospital-born infants undergo circumcision immediately after birth.

In this study, we aimed to evaluate mothers' knowledge and attitudes regarding circumcision, and to raise awareness of common misconceptions that may lead to adverse outcomes in society. Previous studies have largely focused on the medical outcomes of circumcision, while there is limited research addressing mothers' knowledge and attitudes, which are crucial in determining the timing, setting, and safety of the procedure. In this study, we aimed to evaluate mothers' knowledge and attitudes regarding circumcision, and to

raise awareness of common misconceptions that may lead to adverse outcomes in society, while also exploring the influence of sociodemographic factors on these perspectives.

METHODS

The study titled “Assessment of Mothers’ Knowledge and Attitudes Regarding Circumcision” was conducted with the approval of the İstanbul Okmeydanı Training and Research Hospital Ethics Committee (Date: 11.02.2014, Decision No: 161). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

This descriptive, cross-sectional study was conducted among mothers and their children who presented to the pediatric outpatient clinic of a tertiary care hospital. Mothers were selected using a random sampling method. The data were collected through face-to-face interviews using a questionnaire developed based on a literature review and previous similar studies to assess mothers’ knowledge and behaviors regarding circumcision.

The questionnaire included items related to the family’s socioeconomic status, parents’ educational levels, and knowledge, attitudes, and practices concerning circumcision. In cases where the mother had more than one circumcised child, the questions were directed based on the youngest circumcised child.

In this study, male gender was considered an independent variable, while age at circumcision was treated as the dependent variable.

Statistical Analysis

Data analyses were performed using SPSS 15.0 for Windows. Descriptive statistics were presented as numbers and percentages for categorical variables, and as means and standard deviations for numerical variables. For numerical variables that did not follow a normal distribution, intergroup comparisons were conducted using the Mann-Whitney U test. Relationships between categorical variables were evaluated using the chi-square test, and Monte Carlo simulation was applied when necessary. A p-value of less than 0.05 was considered statistically significant.

RESULTS

A total of 540 mothers who had circumcised sons and presented to the pediatric outpatient clinic of a tertiary care hospital for various reasons were included in the study. The mean age of the mothers was 33.3 ± 6.5 years, while the mean age of the fathers was 36.1 ± 6.6 years. The sociodemographic characteristics of the participants are summarized in Table 1.

There were statistically significant differences across geographical regions in terms of circumcision at the appropriate age and knowledge of the correct age for circumcision. These rates were significantly higher among participants born in the Aegean region. Significant differences were also observed between regions regarding the place where circumcision was performed and the person who performed it ($p < 0.001$; $p < 0.001$). It was noteworthy that all circumcisions among participants from the Aegean region were performed

Table 1. Sociodemographic characteristics of the participants

Year		Mean±SD	
Mother's age		33.3±6.5	
Mother's education level		7.1±4.6	
Father's age		36.1±6.6	
Father's education level		8.0±3.9	
		n	Percentage
Place of birth	Black Sea	109	20.1
	Mediterranean	27	5.0
	Aegean	13	2.4
	Marmara	138	25.5
	Eastern Anatolia	110	20.3
	Southeastern Anatolia	68	12.5
	Central Anatolia	83	15.3
District of residence	Şişli	146	26.9
	Kağıthane	87	16.1
	Gaziosmanpaşa	70	12.9
	Beyoğlu	37	6.8
	Levent	30	5.5
	Eyüp	79	14.6
	Others	93	17.2
Mother's occupation	Housewife	418	77.1
	Doctor	22	4.1
	Nurse	21	3.9
	Civil servant	17	3.1
	Worker	41	7.6
	Other	23	4.2
Father's occupation	Unemployed	4	0.7
	Doctor	19	3.5
	Worker	183	33.8
	Civil servant	47	8.7
	Self-employed	265	48.9
	Teacher	13	2.4
	Other	15	2.8
Income level	Below minimum wage	108	19.9
	Minimum wage	309	57.0
	1-2× minimum wage	53	9.8
	≥2× minimum wage	68	12.5
Mean±SD			
Number of children	Girl	0.8±0.8	
	Boy	1.6±0.7	
SD: Standard deviation			

in hospitals and by urologists (100%). Except for the question about penile cancer, statistically significant differences were found between regions in responses to all other questions, with participants born in the Aegean region giving more correct answers. A detailed breakdown of these data is provided in Table 2.

Table 2. Demographic and technical characteristics of circumcised children

		Mean±SD	
Age at circumcision		3.7±2.6	
What is the ideal age for circumcision?		3.0±2.6	
		n	Percentage
Age at circumcision	Newborn	75	13.8
	1-2	135	24.9
	3-6	240	44.3
	>6	92	17.0
Preferred age for circumcision	Newborn	156	28.8
	1-2	81	14.9
	3-6	249	45.9
	>6	56	10.3
Reason for circumcision	Religious	439	81.0
	Medical	359	66.2
	Cultural/traditional	80	14.8
	Other	4	0.7
	UTI	85	15.7
Pre-circumcision anomaly	Phimosis	59	10.9
	Hypospadias	9	1.7
	Crystals	7	1.3
	Hernia	6	1.1
	Stone	2	0.4
	Other	38	7.0
	None	398	73.4
Circumcision technique	With sutures	399	73.9
	Cautery-laser	141	26.1
Reason for timing	Child's fear	232	42.8
	Lack of awareness due to age	208	38.4
	Before starting school	98	18.1
	Financial reasons	32	5.9
	Together with sibling	37	6.8
	Family elders' decision	28	5.2

SD: Standard deviation, UTI: Urinary tract infection

When analyzed by district of residence, there was no statistically significant difference only in the rates of citing religious and traditional reasons for circumcision. In the Levent district, the rate of circumcision for medical reasons was 86.7% (n=26), circumcision at the correct age was 83.3% (n=25), correct knowledge of circumcision age was 93.3% (n=28), hospital-based circumcision was 100% (n=30), the rate of circumcision by a urologist was 70% (n=21), and by a pediatric surgeon was 23.3% (n=7). Mothers in this group also demonstrated high levels of knowledge about complications and gave more correct responses overall. In contrast, mothers from the Gaziosmanpaşa district answered most questions with “I don’t know.” Data on post-circumcision complications and adverse experiences are summarized in [Table 3](#).

Table 3. Post-circumcision complications and mothers' knowledge on the subject

		n	Percentage
Potential postoperative complications of circumcision	Bleeding	412	76.0
	Infection	429	79.2
	Suture dehiscence	66	12.2
	Urinary retention	101	18.6
	Pain	171	31.5
	Physical	43	7.9
	Psychological symptoms	58	10.7
	Other	21	3.9
Observed complication	Bleeding	36	6.6
	Pain	14	2.6
	Infection	10	1.8
	Urinary retention	9	1.7
	Psychological symptoms	8	1.5
	Suture dehiscence	6	1.1
	None	481	88.7

Across occupational groups, statistically significant differences were found in all variables except for the “other” response under known complications. All physician and civil servant mothers (100%) cited medical reasons for circumcision. Among housewives, 64.1% (n=268), nurses 61.9% (n=13), and laborers 63.4% (n=26) cited medical reasons. The rate of circumcision at the appropriate time was 100% among civil servant mothers (n=17), 86.4% among physicians (n=19), 52.2% among housewives (n=218), 56.1% among laborers (n=23), and 19% among nurses (n=4). The rate of correct knowledge of circumcision age was 90.5% among nurses (n=19), 82.4% among civil servants (n=14), 86.4% among physicians (n=19), 49.5% among housewives (n=207), and 36.6% among laborers (n=15). Although nurses had high levels of knowledge, it was notable that they did not have their children circumcised at the appropriate age. All physician, nurse, and civil servant mothers had their children circumcised in hospitals and by medical professionals.

Significant differences were found based on income level in terms of reasons for circumcision and knowledge level. As income increased, the likelihood of citing medical reasons, performing circumcision at the appropriate age, knowing the correct circumcision age, and being aware of possible complications also increased. The rate of hospital-based circumcision and the preference for pediatric surgeons also rose significantly with higher income. Except for the question regarding neonatal circumcision, the rate of correct answers to all other questions increased with income.

According to educational level, statistically significant differences were found in all parameters except for citing religious/traditional reasons, knowing the correct age for circumcision, and awareness of infection risks. Participants who cited medical reasons, correctly identified the appropriate age, and knew about complications had significantly higher years of education ($p<0.001$; for “other” known complications

$p=0.024$). There was a statistically significant association between education level and both the place and the performer of circumcision ($p<0.001$). Mothers whose children underwent mass circumcision had lower educational levels. Mothers who preferred home circumcision had significantly lower education levels than those who chose hospitals; however, there was no statistically significant difference between hospital/polyclinic and home/polyclinic preferences ($p=0.284$; $p=0.094$). The number of years of education was significantly higher among those who preferred urologists and pediatric surgeons compared to those who preferred other healthcare personnel or traditional circumcisers ($p=0.001$; $p<0.001$) (Table 4).

Table 4. Comparative analysis of circumcision characteristics by educational level

			Educational level	
			Mean±SD	p
Circumcision reason	Religious	No	7.8±4.6	0.034
		Yes	7.0±4.6	
	Medical	No	5.6±4.4	<0.001
		Yes	7.9±4.5	
	Cultural	No	7.1±4.6	0.955
		Yes	6.9±4.7	
	Other	No	7.1±4.6	0.051
		Yes	11.5±4.0	
Correct age at circumcision		No	6.6±4.1	0.121
		Yes	7.5±4.9	
Knowledge of correct age		No	5.9±3.8	<0.001
		Yes	8.1±4.9	
Circumcision setting	Home	6.0±4.6 ^{xy}		<0.001
	Hospital	7.9±4.7 ^{xy}		
	Mass circumcision	3.2±3.3 ^{xyz}		
	Outpatient clinic	7.3±2.8 ^y		
Practitioner	Urologist	7.8±4.7		<0.001
	Pediatric surgeon	8.6±4.6		
	General practitioner	6.3±5.0 [*]		
	Health technician	3.6±3.4 ^{xy}		
	Traditional circumciser	5.0±3.8 ^{xy}		
	General surgeon	6.6±2.1		
Complication awareness	Bleeding	No	5.1±4.0	<0.001
		Yes	7.7±4.6	
	Infection	No	7.5±4.1	0.105
		Yes	7.0±4.7	
	Suture dehiscence	No	6.4±4.1	<0.001
		Yes	12.1±5.2	
	Urinary retention	No	6.2±4.0	<0.001
		Yes	11.0±5.1	
	Pain	No	6.1±3.8	<0.001
		Yes	9.3±5.3	
	Physical	No	6.5±4.1	<0.001
		Yes	14.1±4.3	
	Psychological symptoms	No	6.5±4.1	<0.001
		Yes	12.1±5.4	
	Other	No	7.1±4.7	0.024
		Yes	8.5±2.5	

*Different from home #Different from hospital ¥Different from mass circumcision ≠Different from outpatient clinic
^{*}Different from urologist #Different from pediatric surgeon ¥Different from general surgeon
 SD: Standard deviation

Statistically significant differences were found in responses to all questions except for the one asking whether neonatal circumcision should be performed. Participants who responded “I don’t know” to questions regarding urinary tract infections, penile cancer, and sexually transmitted diseases had significantly fewer years of education than those who gave correct answers ($p<0.001$). For the question “Should circumcision be performed for adhesions and phimosis?”, the mean age of those who responded “I don’t know” was significantly lower than that of those who answered correctly or incorrectly ($p<0.001$; $p<0.001$). All comparisons related to the question “Is circumcision essential for fertility?” were statistically significant (correct vs. incorrect $p<0.001$; correct vs. I don’t know $p=0.004$; incorrect vs. I don’t know $p=0.005$) (Table 5).

Table 5. Comparative analysis of responses by educational level

		Educational level	
		Mean±SD	p
UTIs are less common	Correct	8.1±4.5 ^y	<0.001
	Incorrect	7.2±5.3	
	Don’t know	5.0±4.0 [*]	
Penile cancer is less common	Correct	8.5±5.2 ^y	<0.001
	Incorrect	8.4±6.3	
	Don’t know	6.4±4.0 [*]	
Sexually transmitted infections are less common	Correct	8.2±5.0 ^y	<0.001
	Incorrect	8.2±5.5	
	Don’t know	6.2±4.0 [*]	
Circumcision is performed in cases of stenosis or adhesions	Correct	8.3±4.6 ^y	<0.001
	Incorrect	6.9±4.6 ^y	
	Don’t know	4.7±3.6 ^{xy}	
Required for future fertility	Correct	5.1±3.7 ^{xy}	<0.001
	Incorrect	8.8±4.8 ^{xy}	
	Don’t know	6.7±2.7 ^{xy}	
Performed on newborns	Correct	7.2±4.6	0.092
	Incorrect	5.9±4.8	
	Don’t know	8.6±3.0	

*Different from “Correct” #Different from “Incorrect” ¥Different from “I don’t know”
 SD: Standard deviation

DISCUSSION

Circumcision can essentially be performed at any age. However, during the ages of 3 to 6, children are in the phallic stage of development and begin to discover their sexual identity. Children circumcised during this period may fear losing their genitalia, potentially leading to psychological symptoms (e.g., restlessness or anxiety). Therefore, circumcision should be performed before the age of 3 or after the age of 6.^{10,11} The Turkish Urological Association Pediatric Urology Bulletin recommends that circumcision be performed before the age of 2 or after the age of 6, or under general anesthesia if it must be done during the phallic stage.¹² In our study, the mean age of circumcision was 3.7±2.6 years. Among the families, 13.8% had circumcised their child during the neonatal period, 24.9% at 1-2 years, 44.3% at 3-6 years, and 10.3% after the age of 6.

Neonatal circumcision is not widely practiced in Türkiye. In our study, 90% of mothers knew that circumcision could

be performed during the neonatal period; however, the proportion who believed that it was appropriate was low. The majority of those who had their children circumcised in the neonatal period did so based on a physician's recommendation. Although neonatal circumcision is common in other countries, studies from Türkiye report that it is generally not preferred unless medically indicated.^{13,14}

When examining the sociocultural, geographic, economic, and educational backgrounds of the mothers, it was notable that 74.1% in the Mediterranean region, 76.9% in the Aegean region, and 83.3% of mothers living in the Levent district of İstanbul had circumcised their children at the appropriate age. No significant difference was found between income levels and correct timing of circumcision. When occupational groups were analyzed, it was observed that the majority of physician and civil servant mothers had their children circumcised at the appropriate age, while only 52.2% of housewives did so. Interestingly, only 19% of nurse mothers had their children circumcised at the appropriate age. The most frequently cited reasons for choosing a particular age for circumcision were fear (42.8%), lack of awareness (38.4%), medical necessity (21.6%), and completion before starting school (18.1%). These reasons align with the findings of other studies conducted in Türkiye.^{13,14}

Circumcision is commonly performed for religious and traditional reasons, especially among Muslim and Jewish populations. However, due to its medical benefits—such as a lower incidence of penile malignancies and reduced risk of sexually transmitted infections among circumcised males—circumcision is widely practiced globally. In Türkiye, religious and traditional reasons are more dominant than medical ones.¹⁴ In our study, 81% of mothers reported religious reasons, 66.2% reported medical reasons, and 14.8% reported cultural/traditional reasons for circumcision. A study conducted by Özkan et al.¹⁴ in 2017 similarly found that the reasons were primarily religious and traditional. Although most mothers in our study cited religious reasons, it was encouraging that many had started to learn about the medical benefits of circumcision.

When circumcision reasons were compared by birth region, a statistically significant difference was found. Higher rates of medical justification were reported among participants from the Aegean and Marmara regions, and notably, no one from the Aegean region cited tradition as a reason.

It is well known that circumcision performed by trained professionals' results in lower complication rates. In developed countries, where nearly all circumcisions are performed by trained physicians, complication rates range from 0.2% to 5%, typically related to anesthesia.^{15,16} In developing and low-income countries, where circumcision is often performed by untrained individuals and mass circumcisions are common, complication rates can rise to 39% with more severe outcomes.¹⁷ In some studies from Türkiye, the rate of circumcisions performed by trained professionals has been reported as low as 11-12%.^{18,19}

While parents play a major role in the decision to circumcise, the influence of physicians appears limited. In our study, 88.7%

of mothers stated that they experienced no complications post-circumcision; 6.6% reported bleeding, 2.6% pain, 1.8% infection, 1.7% urinary retention, and 1.5% psychological symptoms. These findings are consistent with other studies conducted in Türkiye.^{13,20} Although the low complication rate is reassuring, the results are based on maternal reports and may not be entirely objective. Additionally, complications related to anesthesia used during circumcision were not evaluated in this study.

Regarding knowledge of complications, 79.2% of mothers identified infection, 76% bleeding, 31.5% pain, and 18.6% urinary retention. These findings align with results from previous studies on circumcision complications.^{21,22} The growing accessibility of the internet and the influence of visual and written media on health education may have contributed to increased maternal awareness, which is a positive development. However, it is concerning that many families had no knowledge of the potential psychological symptoms of circumcision performed at inappropriate ages, such as fear, stress, or anxiety.

In our study, although demographic characteristics such as education and income levels were strongly associated with circumcision practices, the central aim was to evaluate maternal knowledge and attitudes. We found that mothers' awareness of potential medical complications, such as bleeding and infection, was relatively high, whereas knowledge regarding psychological symptoms was limited. This finding is consistent with previous studies reporting that families often underestimate the psychological dimension of circumcision.^{11,12} Importantly, the mean age of circumcision in our cohort (3.7 ± 2.6 years) overlaps with the phallic stage of psychosexual development. Circumcision during this period has been linked to regressive behaviors such as enuresis, sleep disturbances, or heightened anxiety, which may reflect the child's fear of genital injury. Although our questionnaire did not directly assess regressive behaviors, future studies could incorporate these outcomes to better understand the psychosocial impact of circumcision timing.

It was determined that 71% of children were not afraid of circumcision. Fear of pain was reported by 24%, fear of bleeding by 12.7%, and fear of genital injury by 6.3% of children.

Information was provided to 35.9% of children prior to circumcision. This information was given by mothers in 29.9% of cases and by fathers in 25.5%. Parents who did not provide information cited concerns that the child might become fearful or was too young to understand. Providing detailed explanations weeks or months in advance may lead to anticipatory anxiety. Not knowing what will happen, having to expose their genitalia, fear of pain, sleep disturbances, restlessness, defiance, and aggression are possible reactions. Therefore, it is recommended that parents offer simple, encouraging explanations shortly before the procedure. It should be explained in age-appropriate language that circumcision is part of cultural and religious traditions and has health benefits. Celebrations, festivities, rewards, and gifts can help reduce anxiety.¹⁰

In our study, maternal knowledge about the benefits of circumcision was also evaluated. A total of 64.4% of mothers correctly identified that circumcised children experience fewer urinary tract infections. Additionally, 58.4% knew that circumcision can be performed in cases of phimosis, adhesions, or narrowing of the foreskin. However, when asked about more complex topics such as the prevention of penile cancer and sexually transmitted diseases, more than half of the mothers responded with “I don’t know.” The literature identifies the medical benefits of circumcision as reduced risk of urinary tract infections, penile cancer, cervical cancer among female partners, and sexually transmitted infections.⁶ As income level increased, the rate of correct responses to these questions also increased.

When asked whether neonatal circumcision is possible, 90% of mothers answered affirmatively. Although this awareness is promising, previous studies from Türkiye indicate that neonatal circumcision is generally avoided unless medically necessary.^{13,14} In our study, mothers stated that neonatal circumcision should only be performed if recommended by a physician or if a medical issue exists.

When asked whether circumcision is necessary for future fertility in boys, 40.8% of mothers responded affirmatively. A total of 50.8% gave incorrect responses, while 8.3% stated they did not know. These findings indicate a lack of knowledge, highlighting the need for further education. The maternal knowledge levels observed in our study are consistent with other studies conducted in Türkiye.^{13,23}

When maternal education levels and occupations were compared with circumcision location and provider, as well as responses to other questions, significant differences were found. Education clearly had a positive effect on determining the appropriate age for circumcision. Our analysis revealed a statistically significant difference between the mothers’ education level and their identification of the ideal circumcision age.

As education level increased, mothers approached circumcision from a more scientific perspective and were better able to understand its medical relevance. Conversely, mothers with lower education levels were more likely to cite cultural or religious reasons for circumcision.²⁴

An analysis of the provider performing the circumcision in relation to maternal education level showed that as education increased, so did the likelihood of choosing a physician for the procedure. All physician, nurse, and civil servant mothers had their children circumcised in a hospital by a medical specialist. This finding suggests that as education increases, individuals are more likely to make informed health decisions.

Limitations

This study has several limitations. First, complications were evaluated only based on maternal self-reports, which may be subject to recall bias or subjective interpretation. Second, the data were collected nearly a decade ago, which may limit its reflection of current practices. However, the large sample size and the inclusion of diverse sociodemographic groups strengthen the value of the findings. Finally, the questionnaire did not include items related to regressive behaviors or broader

psychosocial outcomes, which may have restricted the depth of analysis regarding maternal perspectives.

CONCLUSION

This study revealed that mothers generally have insufficient knowledge regarding circumcision practices, with this deficiency being particularly prominent among those with lower levels of education. A significant portion of participants had their children circumcised between the ages of 3 and 6-considered one of the most psychologically sensitive periods-highlighting the need for greater awareness regarding appropriate timing. On the other hand, the fact that many families chose to have the procedure performed in hospital settings by qualified physicians reflects a growing societal awareness of the importance of safe and medically supervised circumcision. In addition to the prevailing traditional and religious motivations in society, it was observed that mothers are increasingly recognizing the medical benefits of circumcision. Supporting this positive trend and addressing knowledge gaps-especially among mothers with lower education levels-through educational initiatives led by healthcare professionals and trained personnel is of critical importance. Furthermore, in children with appropriate cognitive maturity, the rationale behind circumcision should be communicated in an age-appropriate and supportive manner, integrating medical, traditional, and religious perspectives. Managing this process with attention to the child’s psychological development will contribute to a healthier and more informed approach to circumcision in society.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the İstanbul Okmeydanı Training and Research Hospital Ethics Committee (Date: 11.02.2014, Decision No: 161).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Levine LA, Holzbeierlein JM. Surgery of the Penis and Urethra. In: Wein AJ, Kavoussi LR, Partin AW, Peters CA, eds. Campbell-Walsh-Wein Urology. 12th ed. Vol 1. Elsevier; 2021:992-1022.

2. Zeitler M, Rayala B. Neonatal circumcision. *Prim Care*. 2021;48(4):597-611. doi:10.1016/j.pop.2021.08.002
3. Barthold JS. Anomalies of the penis and urethra. In: Kliegman RM, St. Geme JW, Blum NJ, et al, eds. *Nelson Textbook of Pediatrics*. 22nd ed. Elsevier; 2023. Chapter 559.
4. Karaman Mİ. Her yönüyle sünnet. *Pediatric Üroloji Bülteni*. 2018;1(1):1-4.
5. Dekkers W. Routine (non-religious) neonatal circumcision and bodily integrity: a transatlantic dialogue. *Kennedy Inst Ethics J*. 2009;19(2):125-146. doi:10.1353/ken.0.0279
6. Schoen EJ, Colby CJ, Ray GT. Newborn circumcision decreases incidence and costs of urinary tract infections during the first year of life. *Pediatrics*. 2000;105(4 Pt 1):789-793. doi:10.1542/peds.105.4.789
7. Omole F, Smith W, Carter-Wicker K. Newborn circumcision techniques. *Am Fam Physician*. 2020;101(11):680-685.
8. Yazıcı M, Etensel B, Gürsoy H. Sünnet komplikasyonları. *Adnan Menderes Üniversitesi Tıp Fakültesi Dergisi*. 2003;4(3):5-7.
9. Iacob SI, Feinn RS, Sardi L. Systematic review of complications arising from male circumcision. *BJUI Compass*. 2021;3(2):99-123. doi:10.1002/bco2.123
10. Uprak TK. Recommendation versus current status: retrospective analysis of age to circumcise in a district hospital. *Istanbul Med J*. 2023;24(1):1-6.
11. Yavuz F, Kobya Bulut H. The effect of circumcision age on anxiety and self-perception in boys. *Compr Child Adolesc Nurs*. 2024;47(3):214-223. doi:10.1080/24694193.2024.2389420
12. Silay MS. Male circumcision during phallic period and psychological perspective. *Türk Clin J Urol Spec Topics*. 2018;11(1):45-48.
13. Yılmaz Y, Özsoy SA, Ardahan M. Annelerin sünnet hakkındaki davranış ve bilgi düzeylerinin incelenmesi. *Ege Tıp Dergisi*. 2008;47:93-101.
14. Özkan H. Annelerin sünnet hakkındaki bilgi düzeylerinin belirlenmesi. In: 3. Uluslararası, 4. Ulusal Ebelik Kongresi; May 2017; Ankara, Türkiye.
15. Hutson JM. Male circumcision complications: a systematic review, meta-analysis and meta-regression. *J Urol*. 2021;205(5):1234-1240.
16. Task Force on Circumcision, Blank S, Brady M, et al. American Academy of Pediatrics Task Force on Circumcision. Circumcision policy statement. *Pediatrics*. 2012;130(3):585-586.
17. Appiah Boakye S, Obeng F. Clinical outcomes of circumcisions and prevalence of complications: a five-year retrospective analysis at a teaching hospital in Ghana. *PLOS Glob Public Health*. 2025;5(2):e0003788. doi:10.1371/journal.pgph.0003788
18. Benli E, Koca O. Bingöl ilinde sünnet araştırması. *Yeni Ürol Derg*. 2011;6:22-25.
19. Geçit I, Benli E. Bölgemizdeki sünnet komplikasyonları. *Van Tıp Derg*. 2011;18(18):155-158.
20. Ferhatoglu MF, Kartal A, Gürkan A. Evaluation of male circumcision: retrospective analysis of one hundred and ninety-eight patients. *Cureus*. 2019;11(4):e4555. doi:10.7759/cureus.4555
21. Özkan A, Özorak A, Oruç M. Bin dokuz yüz sünnet olgusunda komplikasyonların retrospektif incelenmesi. *Konuralp Tıp Derg*. 2012;4(1):8-12.
22. Yapanoğlu T, Aksoy Y, Atmaca AF, et al. Bölgemizdeki sünnet komplikasyonları. *Türk Ürol Derg*. 2004;30(4):441-445.
23. Çataklı T, Yazarlı E, Yener F, Bilge YD. Bir hastaneye başvuran annelerin sünnet hakkındaki bilgi düzeyleri. *Erciyes Med J*. 2012;34(3):116-120. doi:10.5152/etd.2012.35
24. Öztürk O, Samut Bülbül S, Durmuşoğlu K. Ebeveynlerin çocuklarda sünnet operasyonuna yönelik bilgi ve tutumlarının değerlendirilmesi. *J Med Clin*. 2024;7(1):11-19.

Evaluation of the MAPH score in predicting acute ischemic stroke severity and major vessel occlusion in the emergency department

Emine Sarcan, Merve Yazla

Department of Emergency Medicine, Ankara Etlik City Hospital, Ankara, Türkiye

Cite this article as: Sarcan E, Yazla M. Evaluation of the MAPH score in predicting acute ischemic stroke severity and major vessel occlusion in the emergency department. *J Health Sci Med.* 2025;8(6):1041-1048.

Received: 30.08.2025

Accepted: 30.09.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to evaluate the utility of the MAPH score, a biomarker combining blood viscosity indicators such as mean platelet volume (MPV), total protein, and hematocrit, in differentiating acute ischemic stroke (AIS) from transient ischemic attack (TIA) and predicting major vessel occlusion in patients presenting to the emergency department with suspected stroke.

Methods: A retrospective analysis was conducted on 226 patients presenting to the emergency department with focal neurological symptoms. Patients were categorized into AIS and TIA groups based on diffusion-weighted magnetic resonance imaging findings. Data collected included demographic characteristics, vital signs, laboratory parameters, and imaging results. MPV, age, hematocrit, and total protein levels were recorded. High-shear rate and low-shear rate were calculated from total protein and hematocrit values. ROC curve analysis was performed to evaluate diagnostic performance.

Results: The ROC analysis demonstrated the diagnostic accuracy of the MAPH score in differentiating AIS from TIA and predicting major vessel occlusion. Additionally, the score showed a significant correlation with NIHSS ≥ 20 , indicating its association with stroke severity.

Conclusion: The MAPH score is a simple and practical tool that aids in distinguishing AIS from TIA and predicting major vessel occlusion, thus improving stroke management in emergency settings.

Keywords: MAPH score, acute ischemic stroke, transient ischemic attack, major vessel occlusion, emergency medicine

INTRODUCTION

Cerebrovascular diseases continue to be a significant cause of mortality and disability in public health. Acute stroke occurs as a result of neuronal damage caused by reduced blood flow to the brain and is a common medical emergency encountered in emergency departments.¹ Globally, approximately 80% of strokes are ischemic, while 20% are hemorrhagic.² Following cardiovascular diseases and cancer, stroke ranks as the third leading cause of death worldwide and is the most prevalent cause of disability.³

Acute ischemic stroke (AIS) occurs due to etiologies such as thrombosis, embolization, and lacunar infarction resulting from small vessel disease. Increased whole blood viscosity (WBV) and the subsequent formation of thrombus are among the primary risk factors for ischemia.⁴ The severity of ischemia may vary depending on the persistence and size of the thrombus, as well as underlying risk factors such as the patient's gender, age, and smoking habits. In some cases, the patient's symptoms may be transient. When symptoms resolve within 24 hours without evidence of tissue infarction on diffusion-weighted magnetic resonance imaging (MRI), the condition is referred to as a transient ischemic attack (TIA).⁵ Factors such as the patient's clinical condition, the size

of the ischemic area, and eligibility for reperfusion strategies necessitate treatment decisions to be made in the emergency department as quickly as possible. The first hours in AIS are critical for preserving brain tissue. While thrombolytic therapies can be administered to eligible patients within the first 4-5 hours of symptom onset, endovascular thrombectomy may be considered in selected patients during extended time windows of up to 24 hours. Clinical criteria are required to grade the severity of ischemic stroke, assess changes during the follow-up process, ensure uniform communication among healthcare professionals and predict prognosis. One of the most commonly used scales, the National Institutes of Health Stroke Scale (NIHSS), evaluates neurological findings by assigning scores and provides valuable information about long-term prognosis. The MAPH score, defined by Abacıoğlu et al.,⁶ incorporates blood viscosity biomarkers such as mean platelet volume (MPV), total protein, hematocrit, and age, and has been used to assess thrombus burden in myocardial infarction (MI) patients. This score was later investigated in the context of thrombus risk in pulmonary thromboembolism (PTE).⁷ Similarly, increased plasma viscosity and alterations in blood rheology can be associated with thrombus formation in ischemic strokes.

Corresponding Author: Emine Sarcan, dr.e.sarcan2010@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

The aim of this study is to evaluate the relationship between major vessel occlusion, AIS, and less critical TIA with the MAPH score in patients presenting to the emergency department with suspected ischemic stroke. Furthermore, this practical score is aimed to contribute to the management protocols for AIS in the emergency department.

METHODS

Study Design and Setting

The study was conducted with approval from the Ankara Etlik City Hospital Scientific Researches Ethics Committee (Date: 25.09.2024, Decision No: AEŞH-BADEK-2024-891). As this was a retrospective file review and did not involve patient identification, informed consent was not required. The study was conducted in accordance with the Declaration of Helsinki. This retrospective study was conducted at the Emergency Medicine Clinic of Ankara Etlik City Hospital between December 1, 2023, and May 1, 2024. The study included male and female patients aged 18 years and older who presented with focal neurological symptoms at least 24 hours after their last known well state. Focal neurological symptoms were defined as the sudden onset of new neurological signs such as asymmetric weakness, sensory changes, dysarthria, facial paralysis, altered mental status, dizziness, and visual impairment. During triage, a critical pathway was activated to facilitate rapid communication with the neurology team for patients suspected of having a stroke. Patients were excluded from the study if they were under 18 years of age, pregnant, had a diagnosis of cerebrovascular disease, pulmonary thromboembolism, or acute coronary syndrome within the past month, were using anticoagulants, or had incomplete laboratory results. A total of 412 patients presenting with suspected acute stroke were evaluated during the specified dates. Stroke was ruled out in 101 of these patients. Additionally, 45 patients who did not meet the inclusion criteria and 40 patients diagnosed with hemorrhagic stroke were excluded from the study. Consequently, a total of 226 patients were included in the study (Figure 1).

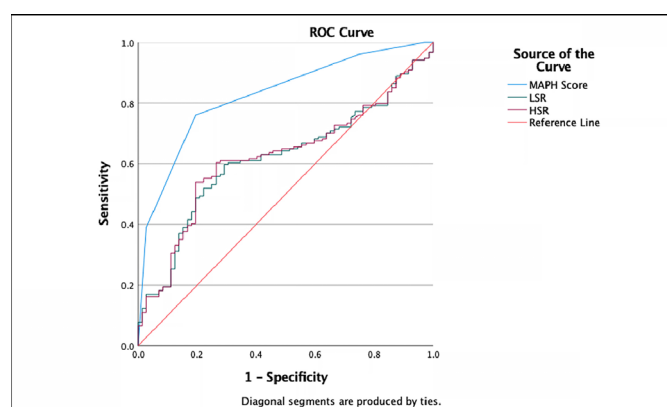


Figure 1. Flow diagram of patient selection and study enrollment

Data Collection and Research Protocol

All patient data were obtained from the hospital database and patient records. Sociodemographic characteristics (age, gender, comorbidities), vital parameters [systolic and diastolic blood pressure (SBP, DBP), body temperature, heart rate

(HR), respiratory rate (RR), peripheral oxygen saturation (SpO₂), Glasgow Coma Scale (GCS), complete blood count parameters [leukocyte, hemoglobin, hematocrit, platelet, MPV], glucose, blood urea nitrogen (BUN), creatinine, total protein (TP), albumin, aspartate aminotransferase (AST), alanine aminotransferase (ALT), sodium, cholesterol, triglycerides, NIHSS, MAPH score, low-shear rate (LSR), high-shear rate (HSR), brain computed tomography (CT), brain computed tomography angiography (CTA), diffusion MRI findings, admission status, and in-hospital mortality data were recorded. All data were reviewed and verified by two emergency medicine specialists. Patients were initially divided into two groups based on diffusion MRI findings: those with pathological findings were classified as the AIS group, while those without pathological findings were classified as the TIA group. The MAPH score, LSR, and HSR values were compared between these two groups. Additionally, in the AIS group, subgroups were formed based on brain CTA findings as patients with and without major vessel occlusion, and the MAPH scores of these subgroups were compared.

Definition

The diagnosis of AIS is supported by radiological imaging and confirmed by the presence of vascular occlusion or detection of diffusion restriction. On the other hand, TIA is defined as a transient episode of neurological dysfunction caused by focal ischemia of the brain, spinal cord, or retina, which resolves rapidly within 24 hours and shows no evidence of tissue infarction on diffusion-weighted MRI.⁵

In AIS, major vessel occlusion was defined as occlusion of the internal carotid artery, middle cerebral artery (M1 and proximal M2 segments), basilar artery, and intracranial vertebral arteries, as identified on brain CTA. Distal M2, A2, and P2 segments were not included in the LVO definition, as these are generally classified as medium vessel occlusions.

The MAPH score consists of four parameters: MPV (M), age (A), total protein (P), and hematocrit (H).⁶ Cut-off values for each parameter of the score were determined using receiver operating characteristic (ROC) curve analysis and the Youden Index. A score of 1 was assigned for values equal to or above the cut-off, and a score of 0 for values below the cut-off. The total MAPH score was calculated by summing the scores of each parameter.

The NIHSS is a scale used to assess neurological status and determine stroke severity.⁸ It consists of an 11-step evaluation system, including parameters such as level of consciousness, eye movements, visual fields, facial paralysis, motor strength (upper and lower extremities), ataxia, sensation, speech, dysarthria, and neglect.

WBV was calculated from Hct and plasma TP at a LSR (0.5 s⁻¹) and a HSR (208 s⁻¹) using the previously validated de Simone formulae.⁹ In our dataset, TP values were originally recorded in g/dL from hospital laboratory records and were converted to g/L (×10) prior to calculation, while Hct was expressed as a percentage. Therefore, the equations were applied using the correct units as follows:

$$\text{HSR: WBV (208 s}^{-1}\text{)} = (0.12 \times \text{HCT}) + 0.17 \times (\text{TP} - 2.07)$$

$$\text{LSR: WBV (0.5 s}^{-1}\text{)} = (1.89 \times \text{HCT}) + 3.76 \times (\text{TP} - 78.42)$$

Statistical Analysis

Descriptive statistics were presented as frequencies (percentages) for categorical variables and as means with standard deviations for numerical variables. The normality assumption for numerical variables was evaluated both analytically and graphically. Comparisons of patient characteristics were performed using the chi-square test for categorical variables and either the Independent Samples T test or the Mann-Whitney U test for numerical variables. The diagnostic ability of laboratory parameters with statistically significant p-values from group comparison tests was assessed using the area under the ROC curve. Cut-off values for potential biomarkers were determined using the Youden Index and diagnostic performance metrics, with sensitivity, specificity, and corresponding 95% confidence intervals, were calculated. Pairwise comparison of ROC curves was performed using the DeLong test. A post-hoc power analysis was also performed for the independent samples T test. Assuming a medium effect size (Cohen's $d=0.5$), an alpha error probability of 0.05, and sample sizes of $n=154$ (group 1) and $n=72$ (group 2), the achieved statistical power ($1-\beta$) was calculated as 93.7%, indicating that the study had sufficient power to test the relevant hypothesis. Statistical analyses were conducted using the Statistical Package for Social Sciences (SPSS, Version 23, Inc., Chicago, IL) software, and a $p<0.05$ was considered statistically significant.

RESULTS

A total of 226 patients were included in the study. Based on clinical findings and diffusion MRI results, 154 patients with acute diffusion restriction were classified as the AIS group, while 72 patients were classified as the TIA group. Sociodemographic data, symptom and examination findings, scores such as NIHSS and GCS, and other findings related to the patients are presented in detail in Table 1. In the AIS group, Hct and MPV values were found to be higher, and this difference was statistically significant.

Based on brain CTA findings, major vessel occlusion was detected in 61 patients (39.6%) in the AIS group. Intravenous thrombolytic therapy with recombinant tissue plasminogen activator (rTPA) was administered to 12 patients (5.3%), and mechanical thrombectomy was performed on 31 patients (13.7%). Among all patients, 44 (19.5%) were discharged from the emergency department, 106 (46.9%) were admitted to the ward, and 76 (33.6%) were transferred to the intensive care unit. In-hospital mortality occurred in 34 patients (19.9%).

In the ROC analysis performed for the MAPH score, the cut-off values were determined according to the Youden index as follows; 11 for MPV, 60 for age, 6.49 for TP, and 43.1 for Hct. The cut-off value for the MAPH score was found to be 3. Comparisons of MAPH, LSR, and HSR values between AIS and TIA patients based on these cut-off values are presented in Table 2.

The ROC analysis demonstrated that the MAPH score, LSR, and HSR could predict the diagnosis of AIS. Table 3 presents all potential cut-off points for the continuous variables that could be used for AIS diagnosis, along with their predictive performance metrics. A comparison of the ROC curves for

the continuous variables revealed that the MAPH score was significantly superior to the others in predicting AIS based on the dataset, followed by LSR and HSR (Figure 2).

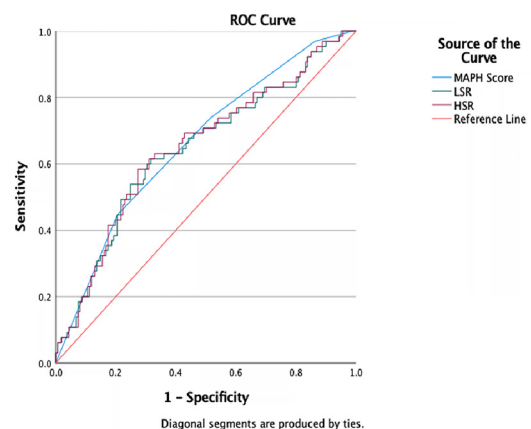


Figure 2. ROC curve of MAPH for patients diagnosed with acute ischemic stroke major vessel occlusion

ROC: Receiver operating characteristic, MAPH: MPV, age, total protein, and hematocrit, LSR: Low shear rate, HSR: High shear rate

Comparisons of MAPH, LSR, and HSR values with the presence of major vessel occlusion in the AIS group are presented in Table 4. ROC analysis demonstrated that the MAPH score, LSR, and HSR could predict the presence of major vessel occlusion in AIS patients. Table 5 details all potential cut-off points for the continuous variables used to determine the presence of major vessel occlusion and the accuracy of these variables in prediction. A comparison of the ROC curves for the continuous variables revealed that MAPH, LSR, and HSR were statistically significant in predicting the presence of major vessel occlusion based on the dataset (Figure 3). According to the DeLong test, there was no statistically significant difference between the MAPH score and LSR ($p=0.657$) or between the MAPH score and HSR ($p=0.777$) in predicting major vessel occlusion. However, the comparison between LSR and HSR demonstrated a statistically significant difference ($p=0.030$), suggesting that these two viscosity-based measures may differ slightly in their diagnostic performance. Overall, the MAPH score showed comparable discriminative ability to both LSR and HSR in this context (Table 6).

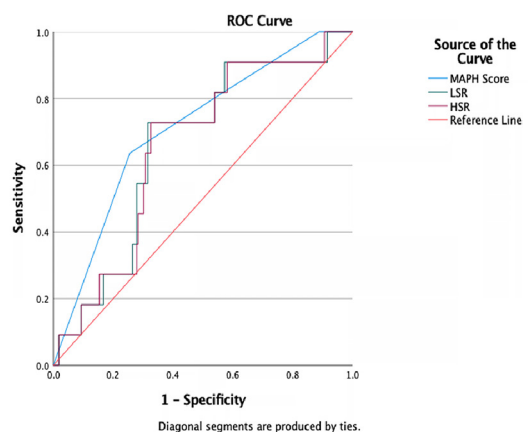


Figure 3. ROC analysis curves of MAPH, LSR and HSR for major vessel occlusion

ROC: Receiver operating characteristic, MAPH: MPV, age, total protein, and hematocrit, LSR: Low shear rate, HSR: High shear rate

Table 1. Sociodemographic data and laboratory values

Total (n=226)		Acute ischemic stroke (n=154)	Transient ischemic attack (n=72)	p-value
Age-mean (±SD)		73±11	72±12	0.751
Gender (n, %)	Female	84 (54%)	39 (54%)	0.958
	Male	70 (46%)	33 (46%)	
Vital signs (mean±SD)				
Systolic blood pressure (mmHg)		169±26	153±27	<0.001*
Diastolic blood pressure (mmHg)		95±14	88±15	0.002*
Heart rate (beats/min)		84±48	80±13	0.955
Oxygen saturation (SpO ₂ , %)		93±3	94±3	0.026*
Body temperature (°C)		36.7±0.4	36.7±0.4	0.618
Respiratory rate (breaths/min)		17±4	16±3	0.096
NIHSS-median (IQR)		6 (3-12)	2 (1-4)	<0.001*
GCS scores (median, IQR)				
Eye response		4 (4-4)	4 (4-4)	0.039*
Verbal response		5 (4-5)	5 (5-5)	0.013*
Motor response		6 (5-6)	6 (6-6)	<0.001*
Comorbidities (n, %)				
Hypertension		140 (90.3%)	60 (84.5%)	0.203
Diabetes mellitus		58 (37.4%)	30 (42.3%)	0.489
Atrial fibrillation		73 (47.4%)	23 (31.9%)	0.028*
Smoking habit		49 (31.6%)	12 (16.9%)	0.021*
Laboratory findings				
White blood cell count (10 ³ /μL)		9.7±3.7	8.6±2.9	0.054
Hemoglobin (g/dl)		13±2.3	12.8±1.7	0.602
Hematocrit (%)		41±6	39±4	0.001*
Platelet count (10 ³ /μL)		240±79	265±156	0.254
Mean platelet volume (fL)		11±0.8	10.4±0.7	<0.001*
Glucose (mg/dl)		150±56	136±47	0.101
Total protein (g/L)		60.5±0.6	60.5±0.4	0.580
Albumin (g/L)		41.5±30	38.8±4	0.816
Blood urea nitrogen (mg/dl)		46±30	40±16	0.437
Creatinine (mg/dl)		1.1±0.9	0.9±0.3	0.364
Total cholesterol (mg/dl)		180±52	169±52	0.077
Low-density lipoprotein (mg/dl)		115±43	109±41	0.156
High-density lipoprotein (mg/dl)		48±15	48±13	0.727
Triglycerides (mg/dl)		129±73	123±61	0.779
*: Chi-square test Mann-Whitney U/T test, SD: Standard deviation, min: Minimum, NIHSS: National Institutes of Health Stroke Scale, IQR: Inter quartile range, GCS: Glasgow Coma Scale				

*: Chi-square test Mann-Whitney U/T test, SD: Standard deviation, min: Minimum, NIHSS: National Institutes of Health Stroke Scale, IQR: Inter quartile range, GCS: Glasgow Coma Scale

Table 2. Comparison of MAPH, LSR, and HSR between acute ischemic stroke and transient ischemic attack patients

Parameters	Acute ischemic stroke	Transient ischemic attack	p-value
MAPH-MPV	0.77 $\pm$ 0.4	0.19 $\pm$ 0.3	<0.001*
MAPH-age	0.90 $\pm$ 0.2	0.76 $\pm$ 0.4	0.005*
MAPH-protein	0.81 $\pm$ 0.3	0.81 $\pm$ 0.3	0.995
MAPH-hematocrit	0.64 $\pm$ 0.4	0.18 $\pm$ 0.3	<0.001*
MAPH	3.11 $\pm$ 0.8	1.94 $\pm$ 0.7	<0.001*
LSR	25 $\pm$ 12	22 $\pm$ 9	0.005*
HSR	4 $\pm$ 0.7	3.8 $\pm$ 0.5	0.003*

*: Mann-Whitney U/T test, MAPH: MPV, age, total protein, and hematocrit, MPV: Mean platelet volume, LSR: Low shear rate, HSR: High shear rate

In the ROC analysis, the MAPH score was compared with NIHSS and showed a significant correlation with patients having NIHSS $\geq$ 20 ($p=0.018$). A comparison of the ROC curves for the continuous variables demonstrated that the MAPH score was statistically significant in predicting NIHSS $\geq$ 20 based on the dataset (Figure 4).

Table 7 presents in-hospital mortality rates stratified by diagnosis groups and stroke severity (NIHSS). Mortality was 26.6% among patients with AIS and 5.6% among those with TIA. Chi-square analysis demonstrated a statistically significant association between diagnosis group and mortality (Pearson $\chi^2=13.656$, $df=1$, $p<0.001$). When stratified by stroke severity, patients with NIHSS $\geq$ 20 had a mortality rate of

Table 3. ROC analysis of MAPH, LSR, and HSR in acute ischemic stroke and transient ischemic attack patients

Parameters	AUC	Cut-off	Sensitivity (%)	Specificity (%)	PPV, %	NPV, %	95% CI (lower)	95% CI (upper)
MAPH	0.825	3	76	81	89	61	0.769	0.881
LSR	0.617	25.9	60	69	81	45	0.543	0.691
HSR	0.621	4.2	54	80	85	45	0.547	0.695

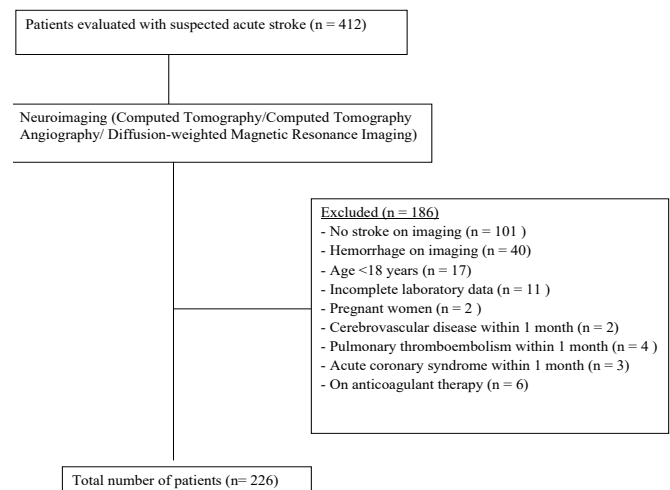
ROC: Receiver operating characteristic, MAPH: MPV, age, total protein, and hematocrit, MPV: Mean platelet volume, LSR: Low shear rate, HSR: High shear rate, AUC: Area under the curve, PPV: Positive predictive value, NPV: Negative predictive value, CI: Confidence interval

Table 4. Comparison of MAPH, LSR, and HSR in patients with and without major vessel occlusion

Parameters	Major vessel occlusion (+)	Major vessel occlusion (-)	p-value
MAPH-MPV	0.80±0.4	0.50±0.5	<0.001*
MAPH-age	0.85±0.3	0.86±0.3	0.738
MAPH-protein	0.78±0.4	0.81±0.3	0.618
MAPH-hematocrit	0.72±0.4	0.40±0.4	<0.001*
MAPH	3.15±0.8	2.5±0.9	<0.001*
LSR	27±11	22±11	<0.001*
HSR	4±0.7	3.8±0.7	<0.001*

*Mann-Whitney U/T test, MAPH: MPV, age, total protein, and hematocrit, MPV: Mean platelet volume, LSR: Low shear rate, HSR: High shear rate

72.7%, compared with 17.2% in those with NIHSS <20. This difference was also statistically significant (Pearson $\chi^2=20.226$, df=1, $p<0.001$).

**Figure 4. ROC analysis of MAPH, LSR and HSR for NIHSS ≥ 20**

ROC: Receiver operating characteristic, NIHSS: National Institutes of Health Stroke Scale, MAPH: MPV, age, total protein, and hematocrit, LSR: Low shear rate, HSR: High shear rate

Table 5. ROC analysis of MAPH, LSR, and HSR in patients with major vessel occlusion

Parameters	AUC	Cut-off	Sensitivity (%)	Specificity (%)	PPV, %	NPV, %	95% CI (lower)	95% CI (upper)
MAPH	0.661	4	45	80	47	78	0.584	0.739
LSR	0.644	27.8	61	68	44	81	0.563	0.726
HSR	0.650	4.25	58	72	46	81	0.569	0.731

ROC: Receiver operating characteristic, MAPH: MPV, age, total protein, and hematocrit, MPV: Mean platelet volume, LSR: Low shear rate, HSR: High shear rate, AUC: Area under the curve, PPV: Positive predictive value, NPV: Negative predictive value, CI: Confidence interval,

Table 6. Pairwise comparisons of ROC curves for major vessel occlusion (DeLong test)

Test pair	z	p (2-tailed)	AUC difference	SE (difference)	95% CI (lower)	95% CI (upper)
MAPH-LSR	0.443	0.657	0.017	0.279	-0.059	0.093
MAPH-HSR	0.284	0.777	0.011	0.279	-0.065	0.087
LSR-HSR	-2.174	0.030	-0.006	0.283	-0.012	-0.001

ROC: Receiver operating characteristic, AUC: Area under the curve, SE: Standard error, CI: Confidence interval, MAPH: MPV, age, total protein, and hematocrit, MPV: Mean platelet volume, LSR: Low shear rate, HSR: High shear rate

Table 7. Mortality stratified by diagnosis and stroke severity (NIHSS)

Mortality	AIS (n=154)	TIA (n=72)	Total (n=226)	p-value*
Survivors	113 (73.4%)	68 (94.4%)	181 (80.1%)	<0.001
Non-survivors	41 (26.6%)	4 (5.6%)	45 (19.9%)	
Total	154 (100%)	72 (100%)	226 (100%)	
Mortality	NIHSS <20 (n=215)	NIHSS ≥ 20 (n=11)	Total (n=226)	p-value
Survivors	178 (82.8%)	3 (27.3%)	181 (80.1%)	<0.001
Non-survivors	37 (17.2%)	8 (72.7%)	45 (19.9%)	
Total	215 (100%)	11 (100%)	226 (100%)	

*: Chi-square test (Pearson χ^2), NIHSS: National Institutes of Health Stroke Scale, AIS: Acute ischemic stroke, TIA: Transient ischemic attack

DISCUSSION

This study demonstrated that the MAPH score exhibited significant performance in differentiating AIS from TIA in patients presenting with suspected stroke and provided high diagnostic accuracy in predicting major vessel occlusion associated with disease severity. By combining parameters such as Htc and TP, which affect blood viscosity, with factors like MPV, reflecting systemic inflammation, and age, the MAPH score can be considered a biomarker aligned with the mechanisms involved in thrombus formation in AIS pathophysiology.

Platelets play an active role in the mechanism of AIS development. The production of thromboxane A2 by activated platelets and the surface presence of glycoprotein IIb/IIIa receptors constitute the primary mechanism of thrombosis formation. This mechanism explains how aspirin and clopidogrel provide antiplatelet effects by inhibiting thromboxane A2 synthesis and glycoprotein IIb/IIIa receptor activity, offering a significant therapeutic approach to reducing the risk of stroke. Although platelet aggregation is a specific method for evaluating platelet function, parameters such as MPV and Platelet Distribution Width also play an important role in the pathophysiology of AIS.¹⁰ MPV is a measure of platelet size and is considered a key indicator of platelet activity.¹¹ In our study, the ROC analysis identified an MPV cut-off value of 11 fL, determined based on the Youden index, which was significantly higher in the AIS group compared to the TIA group and in patients with major vessel occlusion compared to those without. This finding suggests that increased MPV values may be associated with thrombus burden. In another study, MPV was shown to be an independent predictor of stroke risk in 3,134 patients with cerebrovascular disease, and a positive association was observed between MPV and stroke risk in patients with a history of stroke or TIA.¹² Similarly, a meta-analysis conducted by Sadeghi et al.¹³ also identified higher MPV values in AIS patients.

In this study, the ROC analysis identified a hematocrit cut-off value of 43.1%, determined based on the Youden index, which was found to be higher in the AIS group compared to the TIA group and significantly higher in patients with major vessel occlusion compared to those without occlusion. One of the fundamental mechanisms of ischemic stroke is arterial thrombus formation. Increased WBV is a known risk factor for thrombosis, and elevated Htc levels are generally associated with increased WBV.¹⁴ While viscosity increases exponentially with Htc in major vessels, it shows a linear increase in small vessels.¹⁵ In this context, the observed elevation in hematocrit in the AIS group may be associated with the greater thrombus burden in this group compared to the TIA group. Studies on the MAPH score have shown that Htc levels are elevated in patients who have experienced myocardial infarction and have a high thrombus grade.^{6,16}

Total protein, consisting of albumin, fibrinogen, and globulins, is one of the most important serum components regulating plasma viscosity.¹⁷ In one study, higher serum protein levels were found in patients with acute stroke who had a higher degree of spontaneous echo contrast, indicating an increased risk of cardioembolic events.¹⁸ However, in our

study, no significant difference was found between the AIS and TIA groups regarding the determined cut-off value. This may be due to the characteristics of the patient population included in our study and the inability to clearly distinguish factors such as inflammation, hydration status, or nutritional deficiencies that could affect protein levels.

It is well known that AIS is more common, particularly in individuals aged 65 and older, and that mortality and morbidity increase with age. In our study, the >60 years cut-off value determined for predicting AIS was found to be higher in the AIS group compared to the TIA group. This finding is consistent with the literature, which identifies increasing age as a risk factor for thrombus formation and thromboembolic events.¹⁹

Changes in WBV are thought to play an important role in the development of ischemic stroke.²⁰ High WBV in HSR and LSR has been reported to be associated with apical thrombus, stent thrombosis, and poor outcomes in patients with ST-elevation myocardial infarction undergoing percutaneous coronary intervention, as well as an increased risk.²¹ Additionally, elevated WBV has been shown to be a risk factor for both primary and secondary stroke.¹⁴ High WBV at both shear rates has been associated with a high thrombus burden and has been identified as an independent predictor of high thrombus burden in patients with non-ST-elevation myocardial infarction.²² Simone et al.⁹ formulated simple equations using Htc and TP levels to determine WBV at various shear rates. In this study, the comparison of LSR, HSR, and MAPH scores between the TIA and AIS groups revealed that these values were significantly higher in the AIS group. Furthermore, in the AIS group, patients with major vessel occlusion had higher LSR, HSR, and MAPH scores compared to those without occlusion.

One of the most significant findings of the study emerged from the comparison of continuous variables using ROC analyses. The MAPH score, created by incorporating age into the parameters affecting viscosity, such as Htc and TP, demonstrated superior performance compared to known biomarkers like HSR and LSR, as well as its individual components- Htc, age, TP and MPV. This superiority was evident in both the differentiation of AIS and TIA and its association with a high thrombus burden in patients presenting with suspected stroke.

In the ROC analysis, the MAPH score was compared with the NIHSS and showed a significant correlation with patients having NIHSS ≥ 20 . The comparison of ROC curves for continuous variables demonstrated that the MAPH score was statistically significant in predicting high stroke severity at this cutoff. Although different NIHSS severity classifications exist in the literature, in this study NIHSS ≥ 20 was used as the analytic threshold for high severity.^{23,24} This finding is important because therapeutic decisions largely depend on stroke severity and the site of occlusion. Stroke severity may vary depending on the location of the occlusion and thrombus burden, and the MAPH score has the potential to correlate with thrombus load.^{6,16,25} The MAPH score may serve as a complementary tool alongside established assessments such as the NIHSS and neuroimaging modalities in emergency

settings. Early identification of patients at risk for major vessel occlusion is critical for timely reperfusion therapies. Incorporating the MAPH score into initial evaluation protocols could help prioritize patients for advanced imaging and interventions, particularly in resource-limited environments where immediate neuroimaging is not always available.

Limitations

This study has several limitations. One of them is the retrospective design and single-center nature of the study, which limits the generalizability of the results to different patient populations and geographic regions. Additionally, as the MAPH score is a novel biomarker, its validation in larger patient groups and various clinical settings is necessary. Lastly, the study did not thoroughly evaluate factors such as inflammation, hydration status, and nutritional deficiencies, which may influence total protein levels and could limit a comprehensive interpretation of this parameter. Therefore, carefully designed, multicenter prospective studies that also take these factors into account are warranted to strengthen the external validity and generalizability of the MAPH score.

CONCLUSION

This study demonstrated that the MAPH score, a novel indicator of WBV, provides high diagnostic accuracy in both differentiating between AIS and TIA groups and identifying major vessel occlusion. These findings suggest that the MAPH score, which is simple and easy to calculate, could aid in determining disease severity and contribute to patient management in individuals presenting to the emergency department with suspected stroke.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Ankara Etlik City Hospital Scientific Researches Ethics Committee (Date: 25.09.2024, Decision No: AEŞH-BADEK-2024-891).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Walter K. What is acute ischemic stroke? *JAMA*. 2022;327(9):885. doi:10.1001/jama.2022.1420
- Martin SS, Aday AW, Almarzooq ZI, et al. 2024 heart disease and stroke statistics: a report of us and global data from the american heart association. *Circulation*. 2024;149(8):e347-e913. doi:10.1161/CIR.0000000000001209
- Erdem AB, Çavuş UY, Ceylan A, Kaya U, Büyükcım F, Kavalcı C. The prognostic value of leukocyte in cerebrovascular diseases. *Sci Healthcare*. 2020;22(2):61-65. doi:10.34689/SH.2020.22.2.006
- Kang J, Oh JS, Kim BJ, et al. High blood viscosity in acute ischemic stroke. *Front Neurol*. 2023;14:1320773. doi:10.3389/fneur.2023.1320773
- Amin HP, Madsen TE, Bravata DM, et al. Diagnosis, workup, risk reduction of transient ischemic attack in the emergency department setting: a scientific statement from the American Heart Association. *Stroke*. 2023;54(3):e109-e121. doi:10.1161/STR.0000000000000418
- Abacioglu OO, Yildirim A, Karadeniz M, et al. A new score for determining thrombus burden in STEMI patients: the MAPH score. *Clin Appl Thromb Hemost*. 2022;28:10760296211073767. doi:10.1177/10760296211073767
- Akhan O, Boz M, Guzel T, Kis M. Discrimination of the acute pulmonary embolism subtypes based on the novel MAPH score. *J Thromb Thrombolysis*. 2024;57(4):683-690. doi:10.1007/s11239-024-02952-9
- Brott T, Adams HP Jr, Olinger CP, et al. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*. 1989;20(7):864-870. doi:10.1161/01.str.20.7.864
- de Simone G, Devereux RB, Chien S, Alderman MH, Atlas SA, Laragh JH. Relation of blood viscosity to demographic and physiologic variables and to cardiovascular risk factors in apparently normal adults. *Circulation*. 1990;81(1):107-117. doi:10.1161/01.cir.81.1.107
- Zheng YY, Wang L, Shi Q. Mean platelet volume (MPV) and platelet distribution width (PDW) predict clinical outcome of acute ischemic stroke: a systematic review and meta-analysis. *J Clin Neurosci*. 2022;101:221-227. doi:10.1016/j.jocn.2022.05.019
- Farah R, Samra N. Mean platelets volume and neutrophil to lymphocyte ratio as predictors of stroke. *J Clin Lab Anal*. 2018;32(1):e22189. doi:10.1002/jcla.22189
- Bath P, Algert C, Chapman N, Neal B; PROGRESS Collaborative Group. Association of mean platelet volume with risk of stroke among 3134 individuals with history of cerebrovascular disease. *Stroke*. 2004;35(3):622-626. doi:10.1161/01.STR.0000116105.26237.EC
- Sadeghi F, Kovács S, Zsóri KS, Csiki Z, Bereczky Z, Shemirani AH. Platelet count and mean volume in acute stroke: a systematic review and meta-analysis. *Platelets*. 2020;31(6):731-739. doi:10.1080/09537104.2019.1680826
- Lowe GD, Lee AJ, Rumley A, Price JF, Fowkes FG. Blood viscosity and risk of cardiovascular events: the Edinburgh artery study. *Br J Haematol*. 1997;96(1):168-173. doi:10.1046/j.1365-2141.1997.8532481.x
- Pries AR, Neuhaus D, Gaehetgens P. Blood viscosity in tube flow: dependence on diameter and hematocrit. *Am J Physiol*. 1992;263(6 Pt 2):H1770-H1778. doi:10.1152/ajpheart.1992.263.6.H1770
- Çakmak Karaaslan Ö, Çötelci C, Özilhan MO, et al. The predictive value of MAPH score for determining thrombus burden in patients with non-ST segment elevation myocardial infarction. *Egypt Heart J*. 2022;74(1):60. doi:10.1186/s43044-022-00299-1
- Kwaan HC. Role of plasma proteins in whole blood viscosity: a brief clinical review. *Clin Hemorheol Microcirc*. 2010;44(3):167-176. doi:10.3233/CH-2010-1271
- Briley DP, Giraud GD, Beamer NB, et al. Spontaneous echo contrast and hemorheologic abnormalities in cerebrovascular disease. *Stroke*. 1994;25(8):1564-1569. doi:10.1161/01.str.25.8.1564
- Schouten HJ, Geersing GJ, Koek HL, et al. Diagnostic accuracy of conventional or age adjusted D-dimer cut-off values in older patients with suspected venous thromboembolism: systematic review and meta-analysis. *BMJ*. 2013;346:f2492. doi:10.1136/bmj.f2492
- Bilgiç AB, Akpınar ÇK. Association of whole blood viscosity and severe extracranial carotid artery stenosis. *Glob Emerg Crit Care*. 2023;2(2):67-70. doi:10.4274/globecr.galenos.2023.44154
- Ekizler FA, Cay S, Tak BT, et al. Usefulness of the whole blood viscosity to predict stent thrombosis in ST-elevation myocardial infarction. *Biomark Med*. 2019;13(15):1307-1320. doi:10.2217/bmm-2019-0246

22. Çınar T, Şaylık F, Akbulut T, et al. The association between whole blood viscosity and high thrombus burden in patients with non-ST elevation myocardial infarction. *Kardiol Pol.* 2022;80(4):429-435. doi:10.33963/KP.a2022.0043
23. Koton S, Pike JR, Johansen M, et al. Association of Ischemic Stroke Incidence, Severity, and Recurrence With Dementia in the Atherosclerosis Risk in Communities Cohort Study. *JAMA Neurol.* 2022;79(3):271-280. doi:10.1001/jamaneurol.2021.5080
24. Lyden P. Using the National Institutes of Health Stroke Scale: a cautionary tale. *Stroke.* 2017;48(2):513-519. doi:10.1161/STROKEAHA.116.015434

Diagnosis, timing of procedure, and perinatal outcomes in fetocide cases due to fetal anomalies: a five-year experience from Türkiye

 Ebu Bekir Sıddık Yılmaz,  Erkan Sağlam

Department of Perinatology, Bursa City Hospital, Bursa, Türkiye

Cite this article as: Yılmaz EBS, Sağlam E. Diagnosis, timing of procedure, and perinatal outcomes in fetocide cases due to fetal anomalies: a five-year experience from Türkiye. *J Health Sci Med.* 2025;8(6):1049-1054.

Received: 20.06.2025

Accepted: 09.10.2025

Published: 25.10.2025

ABSTRACT

Aims: To evaluate the etiologies of fetal anomalies and the perinatal outcomes of pregnancies undergoing medically indicated selective fetocide.

Methods: This retrospective observational study analyzed 76 pregnancies that underwent fetocide at the Perinatology Unit of Bursa City Hospital between January 2020 and January 2025. Maternal demographics, types of fetal anomalies, and delivery outcomes were systematically reviewed. In dichorionic twin pregnancies, perinatal outcomes of the surviving co-twin were assessed separately. Fetocide was predominantly performed after 21 weeks of gestation by experienced perinatologists under ultrasound guidance, using intracardiac potassium chloride (KCl) injection to induce fetal asystole.

Results: Central nervous system anomalies were the most common indication for fetocide (39.5%), followed by chromosomal anomalies (19.7%) and cardiovascular anomalies (17.1%). The most frequent diagnoses included anencephaly, spina bifida, trisomy 21, and hypoplastic left heart syndrome. In dichorionic twin pregnancies, the mean gestational age at delivery of the surviving fetus was 32.8 weeks, with a mean birth weight of 2266 g. Among these, 71.4% had a 1-minute Apgar score ≥ 7 and 78.6% had a 5-minute Apgar score ≥ 8 . The neonatal intensive care unit (NICU) admission rate was 9.2%.

Conclusion: Central nervous system anomalies were the most common indication for selective fetocide, followed by chromosomal and cardiovascular anomalies. When performed with appropriate timing and patient selection, selective fetocide in dichorionic twin pregnancies it may be a safe and effective approach for the surviving fetus. This study provides valuable data from Türkiye and serves as a foundation for future multicenter prospective research.

Keywords: Fetocide, selective termination, KCl injection, twin pregnancy outcomes, fetal anomaly

INTRODUCTION

Advancements in high-resolution ultrasonography, the widespread implementation of prenatal screening tests, and the increasing availability of diagnostic procedures have led to more frequent detection of fetal anomalies at different stages of pregnancy worldwide. The diagnosis of a congenital anomaly represents a highly stressful situation for the pregnant woman, her partner, and the healthcare team. The decision to terminate the pregnancy (termination of pregnancy, TOP) involves a complex process influenced by medical, ethical, religious, and emotional factors.¹

Some anomalies are detected only in the later stages of pregnancy, or their prognosis becomes clearer as gestation advances. Consequently, late-term TOP may be recommended in selected cases.^{2,3} In general, TOP due to fetal anomalies is considered only when the condition is incompatible with life or associated with a high risk of irreversible, severe physical or mental disability, and when a definitive diagnosis has been established.⁴

In Türkiye, TOP is legally permitted under three circumstances: (1) elective termination, allowed up to 10 weeks of gestation

upon parental request;⁵ (2) maternal health conditions, where termination is permitted at any gestational age if the pregnancy endangers the mother's life; and (3) serious fetal anomalies associated with high morbidity or mortality.^{5,6}

As gestational age advances, the likelihood of live birth following TOP increases. This creates ethical and practical challenges, particularly when a viable neonate is delivered alive despite attempted termination.⁷ In such cases, decisions regarding neonatal intensive care become relevant, and palliative care protocols must be followed as long as the infant remains alive.⁸ The Royal College of Obstetricians and Gynaecologists (RCOG) in the UK has emphasized that in late-term TOP, it is not ethically acceptable for the neonate to die from causes unrelated to the underlying anomaly. Therefore, preventing live birth and its associated complications is strongly recommended.⁹

To minimize such risks, fetocide is often performed prior to termination, thereby eliminating the possibility of unintended live birth. This approach helps avoid emotionally and ethically challenging situations for both patients and healthcare

Corresponding Author: Ebu Bekir Sıddık Yılmaz, drebukirilmaz71@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

providers.¹⁰ Accordingly, fetocide is recommended in all late-term TOP procedures at or beyond 21 weeks and 6 days of gestation, regardless of indication, to ensure intrauterine fetal death prior to delivery.⁸ In addition, clinicians may choose fetocide in medically indicated late-term TOP for several reasons, including shortening the duration of induction, facilitating fetal autopsy and advanced investigations, preventing out-of-hospital live births, and aligning with patient preferences.^{11,12}

The aim of this study was to retrospectively evaluate cases of selective fetocide performed due to structural or chromosomal fetal anomalies. We analyzed the distribution of anomalies and the associated perinatal outcomes. In dichorionic twin pregnancies, the perinatal outcomes of the surviving co-twin were also examined to provide up-to-date data on the prognosis of this specific subgroup. As studies focusing on perinatal outcomes after selective fetocide in twin pregnancies remain limited, this study contributes valuable data from Türkiye and offers a foundation for future research in this field.

METHODS

This retrospective observational study was conducted to evaluate pregnancies that underwent medically indicated fetocide at the Perinatology Unit of the Department of Obstetrics and Gynecology, Bursa City Hospital, between January 2020 and January 2025. The study was carried out in accordance with the ethical principles of the Declaration of Helsinki and was approved by the Bursa City Hospital Scientific Researches Ethics Committee (Date: 14.05.2025, Decision No: 2025-10/18). As this was a retrospective study based on anonymized patient records, the requirement for individual informed consent was waived by the institutional ethics committee.

This study included pregnancies in which structural or chromosomal/genetic fetal anomalies were diagnosed during the antenatal period and intrauterine fetal demise (fetocide) was performed for medical indications. Fetocide was carried out in both singleton and dichorionic twin pregnancies. In twin pregnancies, only cases where fetocide was applied to one fetus while the pregnancy continued with the co-twin were included. Exclusion criteria were monochorionic twin pregnancies, spontaneous intrauterine fetal demise, selective fetal growth restriction (sFGR), twin-to-twin transfusion syndrome (TTTS), terminations not based on medical indications, and incomplete medical records.

All fetocide procedures were performed under sterile conditions and continuous ultrasound guidance by experienced perinatologists. Prior to the procedure, patients received detailed counseling and provided written informed consent. During the procedure, 1-3 ml of 15% potassium chloride (KCl) was injected intracardially into the fetal heart to induce asystole. Fetal demise was confirmed by the absence of cardiac activity on ultrasound. The majority of procedures were performed after 21 weeks of gestation.

Eligible cases were identified by screening ICD-10 diagnostic and procedural codes within the hospital information system. Patient files, ultrasound reports, delivery records, and neonatal unit notes were reviewed in detail. Data were extracted into a standardized collection form for analysis.

The collected variables included maternal age, gravida, parity, pregnancy type (singleton or dichorionic twin), gestational age at fetocide, fetal sex, anomaly type, and mode of delivery (vaginal or cesarean). In dichorionic twin pregnancies undergoing selective fetocide, perinatal outcomes of the surviving co-twin such as gestational age at delivery, birth weight, 1st- and 5th-minute Apgar scores, and neonatal intensive care unit (NICU) admission were evaluated separately.

Fetal anomalies were systematically categorized into the following groups: central nervous system (CNS), cardiovascular system (CVS), gastrointestinal system (GIS), genitourinary system (GUS), musculoskeletal system (MSS), craniofacial anomalies, pulmonary anomalies, and chromosomal/genetic disorders. Subdiagnoses within each category (e.g., anencephaly, spina bifida, trisomy 21, hypoplastic left heart syndrome) were recorded in detail.

Statistical Analysis

Data analyses were performed using IBM SPSS Statistics for Windows, Version 26 (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as mean±standard deviation (SD), and categorical variables as numbers and percentages. As this was a descriptive study, no comparative statistical analyses were conducted; findings were summarized using frequency and distribution analyses.

For the dichorionic twin subgroup, the 95% confidence interval (CI) for the live-birth proportion was calculated using the exact Clopper-Pearson method. In addition, a sensitivity power analysis for a one-sample binomial test (H_0 : $p=0.95$, two-sided $\alpha\leq 0.05$) was performed to illustrate the detectable departures from 95% with the available sample size ($n=14$). Under H_0 , the practical rejection region corresponds to $\leq 11/14$ live births (one-sided $\alpha\approx 0.03$; two-sided $\alpha\leq 0.05$). The resulting powers were $\sim 35\%$ for $p=0.85$, $\sim 55\%$ for $p=0.80$, $\sim 72\%$ for $p=0.75$, and $\sim 84\%$ for $p=0.70$.

RESULTS

The case inclusion process is summarized in [Figure 1](#). Of the initial 102 pregnancies evaluated for fetocide, 26 were excluded due to monochorionic twinning ($n=10$), intrauterine fetal demise prior to the procedure ($n=6$), and incomplete medical records ($n=10$). The final analysis included 76 pregnancies, consisting of 62 singletons and 14 dichorionic twins.

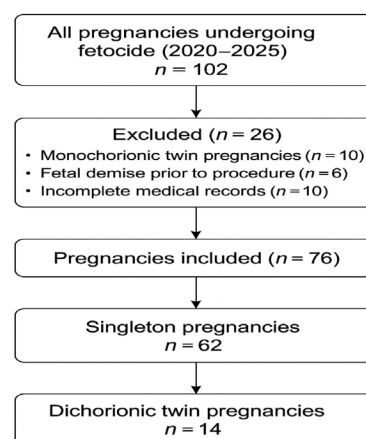


Figure 1. Flowchart of patient selection process

From 102 pregnancies that underwent fetocide evaluation, 26 cases were excluded due to monochorionic twinning ($n=10$), fetal demise before procedure ($n=6$), and incomplete medical records ($n=10$). A total of 76 pregnancies were included in the final analysis, comprising 62 singleton and 14 dichorionic twin pregnancies

The mean maternal age was 32.3 ± 6.8 years, and the mean gestational age at the time of fetocide was 26.7 ± 4.4 weeks. Overall, 19.7% of patients were primigravida and 80.3% were multigravida; 30.3% were nulliparous, while 69.7% had previous births. Singleton pregnancies accounted for 81.6%, and dichorionic twin pregnancies for 18.4%. The fetal sex distribution was 53.9% male and 46.1% female. In dichorionic twin pregnancies resulting in live birth, delivery was vaginal in 84.2% and cesarean in 15.7% (Table 1).

Table 1. Maternal demographic and obstetric characteristics in pregnancies undergoing fetocide

Parameter	Mean $\pm$ SD/n (%)
Maternal age (years)	32.3 $\pm$ 6.8
Gestational age at procedure (weeks)	26.7 $\pm$ 4.4
Gravida (total number of pregnancies)	
Primigravida	15 (19.7)
Multigravida	61 (80.3)
Parity (≥ 24 weeks births)	
Nulliparous	23 (30.3)
Multiparous	53 (69.7)
Type of pregnancy	
Singleton	62 (81.6)
Twin (dichorionic)	14 (18.4)
Fetal sex (fetus undergoing fetocide)	
Male	41 (53.9)
Female	35 (46.1)
Mode of delivery (if live birth)	
Vaginal delivery	64 (84.2)
Cesarean section	12 (15.7)

Maternal demographic and obstetric characteristics of cases undergoing fetocide due to fetal anomalies are summarized. Continuous variables are presented as mean $\pm$ standard deviation (SD), and categorical variables as number (percentage, %). Parameters such as gravida, parity, and mode of delivery are classified based on clinical definitions

Regarding the distribution of anomaly groups, CNS anomalies were the most frequent indication for fetocide (39.5%), followed by chromosomal anomalies (19.7%) and cardiovascular anomalies (17.1%). Other anomalies included those of the genitourinary system (5.3%), gastrointestinal system (5.3%), craniofacial region (3.9%), pulmonary system (3.9%), musculoskeletal system (2.6%), and non-chromosomal genetic disorders (2.6%) (Table 2, Figure 2).

Table 2. Distribution of anomaly groups among fetocide cases

Anomaly group	n	%
Central nervous system	30	39.5
Cardiovascular system	13	17.1
Musculoskeletal system	2	2.6
Chromosomal anomalies	15	19.7
Face and neck anomalies	3	3.9
Gastrointestinal system	4	5.3
Pulmonary system	3	3.9
Genitourinary system	4	5.3
Non-chromosomal genetic diseases	2	2.6
Total	76	100

The distribution of anomaly groups among pregnancies undergoing fetocide is shown. Central nervous system anomalies were the most common indication, followed by chromosomal and cardiovascular anomalies. Percentages were calculated based on the total number of cases (n=76).

Distribution of Anomaly Groups in Fetocide Cases (n=76)

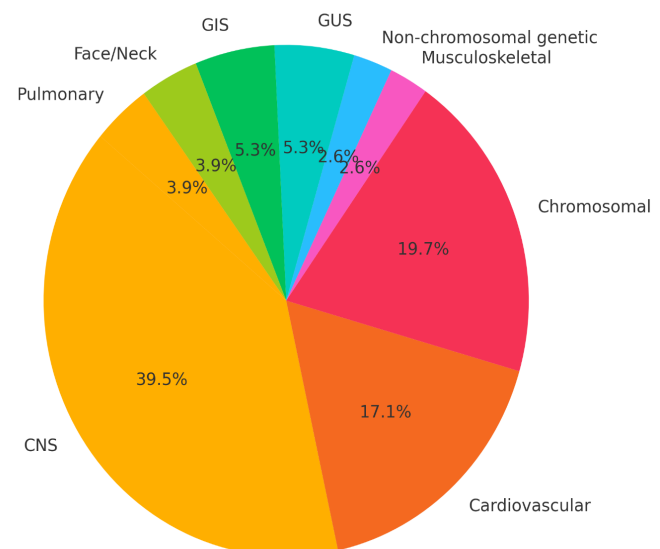


Figure 2. Distribution of anomaly groups among pregnancies undergoing fetocide

The most common indication was central nervous system (CNS) anomalies (39.5%), followed by chromosomal anomalies (19.7%) and cardiovascular anomalies (17.1%). Other anomalies (musculoskeletal, gastrointestinal, genitourinary, pulmonary, face and neck, genetic diseases) were less frequently observed

A more detailed evaluation of anomaly subtypes revealed that the most common diagnoses within the CNS group were anencephaly-acrania-exencephaly (14.5%), spina bifida (10.5%), Dandy-Walker malformation (3.9%), and hydrocephalus (3.9%). Trisomy 21 (11.8%) was the most frequent chromosomal anomaly, followed by isolated cases of trisomy 18, trisomy 13, and triploidy. Among cardiovascular anomalies, hypoplastic left heart syndrome (HLHS) was the most frequent (6.6%), with additional cases of atrioventricular or ventricular septal defects (AVSD/VSD), tetralogy of Fallot, double outlet right ventricle (DORV), and Ebstein anomaly. Other anomalies identified included omphalocele, bilateral renal agenesis, posterior urethral valve (PUV), cystic hygroma, congenital diaphragmatic hernia (CDH), congenital pulmonary airway malformation (CPAM), and lethal skeletal dysplasias (Table 3).

In dichorionic twin pregnancies undergoing selective fetocide, perinatal outcomes of the surviving co-twin were analyzed. The mean gestational age at delivery was 32.8 ± 3.3 weeks, with a mean birth weight of 2266 ± 502 g. Delivery mode was almost equally distributed between vaginal (48.7%) and cesarean (51.3%). The mean Apgar scores were 7.4 ± 1.2 at 1 minute and 8.4 ± 0.9 at 5 minutes. Neonatal intensive care unit (NICU) admission was required in 9.2% of cases (Table 4). All 14 dichorionic twin pregnancies resulted in live birth (100%; 95% CI 76.8-100 by exact Clopper-Pearson). A sensitivity power analysis indicated limited ability to detect modest deviations from 95% (power ~35% for 85% and ~55% for 80%), but adequate power for larger differences (e.g., ~84% for 70%).

The relationship between gestational age at fetocide and delivery week of the surviving co-twin is illustrated in Figure 3. No significant linear correlation was observed, indicating that the timing of fetocide did not substantially influence the gestational age at delivery.

Table 3. Distribution of fetal anomaly subtypes in pregnancies undergoing fetocide

System/subdiagnosis	n	%
Central nervous system	30	39.4
Anencephaly/acrania/exencephaly	11	14.5
Spina bifida	8	10.5
Dandy-Walker malformation	3	3.9
Hydrocephalus	3	3.9
Holoprosencephaly	2	2.6
Agenesis of corpus callosum	3	3.9
Chromosomal anomalies	15	19.7
Trisomy 21	9	11.8
Trisomy 18	3	3.9
Trisomy 13	2	2.6
Triploidy	1	1.3
Cardiovascular anomalies	13	17.1
Hypoplastic left heart syndrome	5	6.6
AVSD/VSD	4	5.3
Tetralogy of fallot	2	2.6
Double outlet right ventricle	1	1.3
Ebstein anomaly	1	1.3
Face and neck anomalies	3	3.9
Cystic hygroma	3	3.9
Gastrointestinal system	4	5.3
Omphalocele	3	3.9
Esophageal atresia	1	1.3
Genitourinary system	4	5.3
Bilateral renal agenesis	2	2.6
Polycystic kidney disease	1	1.3
Posterior urethral valve	1	1.3
Genetic diseases	2	2.6
DiGeorge syndrome	2	2.6
Pulmonary system	3	3.9
Congenital diaphragmatic hernia	2	2.6
CPAM	1	1.3
Musculoskeletal system	2	2.6
Lethal skeletal dysplasia	2	2.6
Total	76	100.0

The distribution of fetal anomaly subtypes by system in pregnancies undergoing fetocide. The most common diagnoses were central nervous system, chromosomal, and cardiovascular anomalies. Percentages were calculated based on a total of 76 cases.

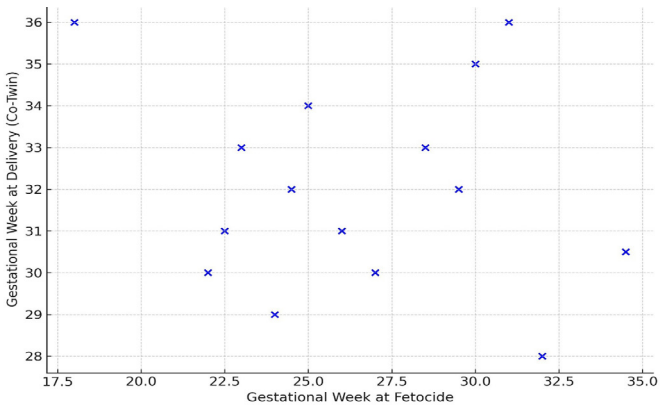


Figure 3. Scatter plot showing the relationship between the gestational week at which selective fetocide was performed and the delivery week of the surviving co-twin in dichorionic twin pregnancies. No significant linear trend was observed, indicating that the timing of fetocide did not markedly affect the co-twin’s gestational age at delivery

DISCUSSION

Termination of pregnancy is not only a medical intervention but also a process shaped by ethical, cultural, and legal factors. In Türkiye, elective termination is permitted up to the 10th gestational week, while after this period it is allowed only in cases of severe fetal anomalies or life-threatening maternal conditions, with the approval of two specialists.^{5,6} Therefore, fetocide procedures in pregnancies with fetal anomalies require careful planning and multidisciplinary counseling.

Congenital structural anomalies or genetic disorders complicate approximately 3-5% of pregnancies.¹³ and constitute one of the main reasons for termination following prenatal diagnosis. CNS anomalies are consistently reported as the most frequent indication,¹⁴⁻¹⁶ a finding confirmed by our study, where they accounted for 39.5% of all fetocide cases.

Advances in ultrasonography, serum biomarkers, and cell-free fetal DNA testing have enabled earlier detection of chromosomal anomalies.^{17,18} In our study, 19.7% of fetocide cases were due to chromosomal anomalies and 2.6% to inherited/genetic disorders, rates consistent with national data but slightly lower than the 25-39% reported in international studies.¹⁹⁻²¹ This difference may reflect cultural, religious, or socioeconomic factors.

When gestational age at fetocide was compared by anomaly type, earlier terminations were seen in genetic disorders (21.5±2.1 weeks) and chromosomal anomalies (25.1±3.9 weeks). Structural anomalies were generally terminated later: 27.2±3.6 weeks for CNS anomalies, 27.1±4.1 weeks for cardiovascular anomalies, and 28.9±2.9 weeks for gastrointestinal anomalies. These findings indicate that screening methods are more effective in detecting genetic and chromosomal anomalies earlier, while structural anomalies often become apparent at later stages. Similar results have been reported previously.^{16,21}

Advances in prenatal cardiology, including nuchal translucency measurement, tricuspid regurgitation assessment, and ductus venosus Doppler, allow earlier diagnosis of cardiac anomalies.²² While Corbacioğlu et al.¹⁵ reported a 1.7% fetocide rate due to cardiac anomalies, our study found 17.1%, likely reflecting the increased use and sensitivity of fetal echocardiography.

Genetic disorders remain largely untreatable during pregnancy, prompting wider use of genetic testing. In our study, fetocide due to genetic diseases was 2.6%, similar to the 1.4% reported by Özyuncu et al.²³ Most cases had detailed family histories and were referred for advanced evaluation, underscoring the importance of multidisciplinary management.

A distinctive aspect of our study is the evaluation of surviving co-twins after selective fetocide in dichorionic twin pregnancies. The goal is to terminate the anomalous fetus while preserving the healthy co-twin. Reported neonatal loss rates are 1.4% for early procedures (<14 weeks) and 4.1% for later procedures, while cohort studies from China and Denmark demonstrated live birth rates of 95-96% with prematurity risk comparable to controls.²⁴

In our study, all 14 dichorionic twin cases that underwent selective fetocide resulted in live births, yielding a 100%

live birth rate. When considered alongside our findings of a 9.2% NICU admission rate, which should be interpreted with caution due to the lack of a control group or comparative population data limiting its clinical significance, a mean gestational age at delivery of 32+6 weeks, adequate birth weights, and high Apgar scores, this supports the view that selective fetocide may represent a feasible intervention with generally favorable short-term outcomes when appropriately timed and correctly indicated, although larger prospective studies are required to confirm its long-term safety and effectiveness. These outcomes are comparable with international cohort studies. For example, Kristensen et al.²⁴ reported NICU admission rates of approximately 10% and live birth rates exceeding 95% following selective fetocide in dichorionic twins. Similarly, large European and Asian series have demonstrated Apgar scores and survival outcomes consistent with our findings, supporting the external validity of our results. Nonetheless, further studies are needed to assess long-term neurodevelopmental outcomes in surviving co-twins. However, the absence of a significant correlation between the timing of fetocide and the delivery week of the surviving co-twin (**Figure 3**) should be interpreted with caution, as the small sample size carries a considerable risk of a type II error.

A total of 84.2% of pregnancy terminations were performed via vaginal delivery, while the cesarean delivery rate was 15.7%. Hysterotomy was required only in cases with a history of uterine surgery or those at advanced gestational ages in the second trimester. This approach is consistent with clinical recommendations aimed at minimizing maternal complications reported in the literature.²⁵ Therefore, termination at the earliest feasible gestational age in pregnancies affected by severe anomalies may reduce maternal morbidity and alleviate both ethical and psychological burden.

Limitations

This study has several limitations. First, the retrospective design may have resulted in missing or inconsistently recorded clinical variables. Moreover, we were unable to conduct long-term neurodevelopmental or pediatric follow-up of the surviving fetuses. The lack of systematic long-term neurodevelopmental evaluation of co-twins represents an important limitation of our study. Additionally, our study included only dichorionic twin pregnancies and thus cannot be generalized to outcomes of selective fetocide in monochorionic twin pregnancies. Although we observed a 100% live-birth rate in the dichorionic subgroup, the exact 95% CI was wide (76.8-100), reflecting imprecision due to small sample size. Consistently, our sensitivity power analysis indicated limited power to detect modest deviations from 95%, underscoring the need for larger multicenter cohorts.

However, the strengths of this study include its five-year data span, systematic analysis of fetal anomaly subtypes, and the evaluation of perinatal outcomes in dichorionic twins-a group for which data from Türkiye are limited. These features provide original and valuable contributions to the relevant literature.

CONCLUSION

This study demonstrated that central nervous system anomalies were the most common indication for fetocide, followed by chromosomal and cardiovascular system anomalies. Severe CNS pathologies such as anencephaly and spina bifida represented the leading causes of pregnancy termination after prenatal diagnosis. In dichorionic twin pregnancies, perinatal outcomes of the surviving co-twin after selective fetocide were generally favorable. Gestational age at delivery and birth weight were within acceptable ranges, Apgar scores were reassuring, and the need for neonatal intensive care was low. These results suggest that, when performed with appropriate case selection and under expert perinatal care, selective fetocide may be considered a feasible option with generally favorable short-term outcomes for the healthy co-twin. However, definitive conclusions about long-term safety and effectiveness require larger multicenter prospective studies. Our findings provide one of the few datasets on fetocide from Türkiye and contribute to the international literature. They may serve as a foundation for future prospective, multicenter studies aimed at refining patient selection, procedural timing, and long-term neonatal outcomes.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Bursa City Hospital Scientific Researches Ethics Committee (Date: 14.05.2025, Decision No: 2025-10/18).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Acknowledgments

The author gratefully acknowledges anonymous support in statistical analysis and English language editing.

REFERENCES

1. Phadke SR, Agarwal M, Aggarwal S. Late termination of pregnancy for fetal abnormalities: the perspective of Indian lay persons and medical practitioners. *Prenat Diagn*. 2011;31(13):1286-1291. doi:10.1002/pd.2887
2. Rydberg C, Tunón K. Detection of fetal abnormalities by second-trimester ultrasound screening in a non-selected population. *Acta Obstet Gynecol Scand*. 2017;96(2):176-182. doi:10.1111/aogs.13125

3. Pasquini L, Pontello V, Kumar S. Intracardiac injection of potassium chloride as method for fetocide: experience from a single UK tertiary centre. *BJOG*. 2008;115(4):528-531. doi:10.1111/j.1471-0528.2007.01646.x
4. Breeze AC, Lees CC, Kumar A, Missfelder-Lobos HH, Murdoch EM. Palliative care for prenatally diagnosed lethal fetal abnormality. *Arch Dis Child Fetal Neonatal Ed*. 2007;92(1):F56-F58. doi:10.1136/adc.2006.099598
5. Resmi Gazete. Law No. 2827 on Population Planning. Tertip: 1983;22: 352. [in Turkish]
6. Gursoy A. Abortion in Türkiye: a matter of state, family or individual decision. *Soc Sci Med*. 1996;42(4):531-542. doi:10.1016/0277-9536(95)00165-4
7. Senat MV, Fischer C, Bernard JP, Ville Y. The use of lidocaine for fetocide in late termination of pregnancy. *BJOG*. 2003;110(3):296-300. doi:10.1046/j.1471-0528.2003.02283.x
8. Govender L, Moodley J. Late termination of pregnancy by intracardiac potassium chloride injection: 5 years' experience at a tertiary referral centre. *S Afr Med J*. 2013;103(1):47-51. doi:10.7196/samj.5967
9. Royal College of Obstetricians and Gynaecologists (RCOG). The care of women requesting induced abortion. Guideline No. 7. London: RCOG Press; 2004.
10. Diedrich J, Drey E; Society of Family Planning. Induction of fetal demise before abortion. *Contraception*. 2010;81(6):462-473. doi:10.1016/j.contraception.2010.01.024
11. Sharvit M, Klein Z, Silber M, et al. Intra-amniotic digoxin for fetocide between 21 and 30 weeks of gestation: a prospective study. *BJOG*. 2019; 126(7):885-889. doi:10.1111/1471-0528.15656
12. Tufa TH, Prager S, Lavelanet AF, Kim C. Drugs used to induce fetal demise prior to abortion: a systematic review. *Contracept X*. 2020;2: 100046. doi:10.1016/j.conx.2020.100046
13. Dar P, Shani H, Evans MI. Cell-free DNA: comparison of technologies. *Clin Lab Med*. 2016;36(2):199-211. doi:10.1016/j.cll.2016.01.004
14. Barel O, Vaknin Z, Smorgick N, et al. Fetal abnormalities leading to third trimester abortion: nine-year experience from a single medical center. *Prenat Diagn*. 2009;29(3):223-228. doi:10.1002/pd.2206
15. Corbacioğlu A, Aslan H, Aydın S, et al. Trends in fetal indications for termination of pregnancy between 2002 and 2010 at a tertiary referral centre. *J Turk Ger Gynecol Assoc*. 2012;13(2):85-90. doi:10.5152/jtgga.2012.10
16. Hern WM. Fetal diagnostic indications for second and third trimester outpatient pregnancy termination. *Prenat Diagn*. 2014;34(5):438-444. doi:10.1002/pd.4331
17. Carlson LM, Vora NL. Prenatal diagnosis: screening and diagnostic tools. *Obstet Gynecol Clin North Am*. 2017;44(2):245-256. doi:10.1016/j.ogc.2017.02.004
18. Cuckle H, Maymon R. Development of prenatal screening: a historical overview. *Semin Perinatol*. 2016;40(1):12-22. doi:10.1053/j.semperi.2015.11.002
19. Dommergues M, Mandelbrot L, Mahieu-Caputo D, et al. Termination of pregnancy following prenatal diagnosis in France: how severe are the foetal anomalies? *Prenat Diagn*. 2010;30(6):531-539. doi:10.1002/pd.2501
20. Wylles MP, Tonks AM. Termination of pregnancy for fetal anomaly: a population-based study 1995 to 2004. *BJOG*. 2007;114(5):639-642. doi: 10.1111/j.1471-0528.2007.01311.x
21. Vaknin Z, Ben-Ami I, Reish O, et al. Fetal abnormalities leading to termination of singleton pregnancy: the 7-year experience of a single medical center. *Prenat Diagn*. 2006;26(10):938-943. doi:10.1002/pd.1527
22. Minnella GP, Crupano FM, Syngelaki A, et al. Diagnosis of major heart defects by routine first-trimester ultrasound examination: association with increased nuchal translucency, tricuspid regurgitation and abnormal flow in ductus venosus. *Ultrasound Obstet Gynecol*. 2020; 55(5):637-644. doi:10.1002/uog.20450
23. Ozyuncu O, Orgul G, Tanacan A, et al. Retrospective analysis of indications for termination of pregnancy. *J Obstet Gynaecol*. 2019;39(3): 355-358. doi:10.1080/01443615.2018.1515197
24. Kristensen SE, Ekelund CK, Sandager P, et al. Risks and pregnancy outcome after fetal reduction in dichorionic twin pregnancies: a Danish national retrospective cohort study. *Am J Obstet Gynecol*. 2023;228(5): 590.e1-590.e12. doi:10.1016/j.ajog.2023.01.018
25. Garofalo G, Garofalo A, Sochirca O, et al. Maternal outcomes in first and second trimester termination of pregnancy: which are the risk factors? *J Perinat Med*. 2018;46(4):373-378. doi:10.1515/jpm-2017-0137

Novel GLA mutation in a high-consanguinity region: insights from Fabry disease screening in Türkiye

✉Sibel Ada¹, ✉Şeyda Gül Özcan², ✉Orçun Altunören³, ✉Necmi Eren⁴, ✉Tayfur Toptaş⁵

¹Department of Nephrology, Prof. Dr. Cemil Taşcıoğlu City Hospital, İstanbul, Türkiye

²Division of Nephrology, Department of Internal Medicine, Cerrahpaşa Faculty of Medicine, İstanbul University, İstanbul, Türkiye

³Division of Nephrology, Department of Internal Medicine, Faculty of Medicine, Kahramanmaraş Sütçü İmam University, Kahramanmaraş, Türkiye

⁴Division of Nephrology, Department of Internal Medicine, Faculty of Medicine, Kocaeli University, Kocaeli, Türkiye

⁵Division of Hematology, Department of Internal Medicine, Faculty of Medicine, Marmara University, İstanbul, Türkiye

Cite this article as: Ada S, Özcan ŞG, Altunören O, Eren N, Toptaş T. Novel GLA mutation in a high-consanguinity region: insights from Fabry disease screening in Türkiye. *J Health Sci Med.* 2025;8(6):1055-1059.

Received: 07.08.2025

Accepted: 08.10.2025

Published: 25.10.2025

ABSTRACT

Aims: Fabry disease (FD) is an X-linked lysosomal storage disorder caused by deficient α -galactosidase A (α -Gal A) activity, leading to the accumulation of globotriaosylceramide in multiple organ systems. Its early diagnosis remains a challenge, particularly in high-risk populations. This study aimed to determine the prevalence of FD among hemodialysis patients in Eastern and Southeastern Türkiye, regions with a high rate of consanguineous marriages.

Methods: Between 2015 and 2019, 613 adult patients undergoing maintenance hemodialysis across five provinces were screened for α -Gal A activity using dried blood spot testing. Patients with reduced enzyme levels underwent confirmatory testing through leukocyte enzyme assays and GLA gene sequencing.

Results: Reduced α -Gal A activity was identified in 15.7% of patients. Genetic analysis in 58 individuals revealed no mutations in males. However, two unrelated female patients were found to carry the same GLA gene variant, c.116C>A (p.T39K), which has not been previously reported in gnomAD or ClinVar databases. The overall prevalence of genetically confirmed FD in this cohort was 0.32%.

Conclusion: This is the first regional screening study of FD in Eastern and Southeastern Türkiye. The identification of a novel GLA mutation in a high-consanguinity population underscores the importance of genetic screening in dialysis patients and highlights the need for region-specific diagnostic and counseling strategies.

Keywords: Fabry disease, hemodialysis screening, GLA gene, consanguinity

INTRODUCTION

Fabry disease (FD) is a rare X-linked lysosomal storage disorder caused by mutations in the GLA gene, which encodes the enzyme α -galactosidase A (α -Gal A).¹ Deficiency or absence of α -Gal A activity leads to the systemic accumulation of globotriaosylceramide (Gb3) and its derivatives in lysosomes, ultimately resulting in multi-organ damage.²⁻⁴ Organs most commonly affected include the kidneys, heart, and nervous system.⁵⁻⁷ Renal involvement is a hallmark of FD and can begin early, even in asymptomatic individuals.^{8,9} Gb3 accumulation contributes to podocyte injury, proteinuria, and progressive decline in renal function, often culminating in end-stage renal disease (ESRD).¹⁰

The GLA gene, located on chromosome Xq22, is highly polymorphic with over 1000 variants described to date.¹¹ Most GLA variants are private, and novel mutations continue to be reported, especially in populations with high rates of consanguinity.^{12,13} In these regions, genetic phenomena such as the founder effect and increased homozygosity may

amplify the prevalence and clinical expression of otherwise rare pathogenic alleles.^{14,15}

Eastern and Southeastern Türkiye are known for their elevated rates of consanguineous marriages.^{16,17} Such demographic characteristics increase the prevalence of autosomal recessive and X-linked disorders, leading to regional clustering of genetic diseases. Therefore, these regions represent a high-risk population for Fabry disease screening.³³

Previous studies conducted in Western and Mediterranean regions of Türkiye have reported Fabry disease prevalence among dialysis patients between 0.1% and 0.2%.^{18,19} Comparable studies from Europe and Asia have shown prevalence rates ranging from 0.2% to 0.3% in dialysis cohorts.²⁰⁻²² By contrast, our study revealed a prevalence of 0.32%, which appears slightly higher than both national and international reports, possibly reflecting the genetic characteristics of this region.

Although several international studies have assessed the prevalence of FD in dialysis populations from Türkiye, data

Corresponding Author: Abdulgani Gülyüz, abdulganigulyuz@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

particularly from high-risk regions where consanguinity rate is high, remain scarce.¹⁸⁻²¹ Therefore, this study aims to determine the prevalence of Fabry disease in dialysis populations from Eastern and Southeastern Türkiye, regions characterized by high consanguinity and a unique genetic background, in order to provide insights for future screening and counseling strategies.

METHODS

Ethics

This study was approved by the Kahramanmaraş Sütcü İmam University Faculty of Medicine Scientific Researches Ethics Committee (Date: 27.07.2015, Decision No: 13). All data were anonymized, and the study was conducted in accordance with the Declaration of Helsinki.

Study Design and Participants

We performed a prospective, multicenter, cross-sectional, screening study between October 2015 and February 2019. We recruited adult patients with ESRD undergoing maintenance hemodialysis (HD), from five dialysis centers located in the East and Southeast regions of Türkiye.

Patients under the age of 18 or without confirmed ESRD were excluded. A total of 613 patients were enrolled the study. Among the patients screened, no one had been diagnosed with FD.

Screening Procedure and Clinical Evaluation

Peripheral blood and dried blood spot (DBS) samples were collected from each participant. The first 98 patients were analyzed at Universitätsklinikum Hamburg-Eppendorf, where α-Gal A enzyme activity was measured using a mass spectrometry-based assay and reported in pmol/spot/21 hours. The remaining 515 patients were tested at ARCHIMED Life Science GmbH (Vienna), using a plasma-based fluorometric assay with results reported in μmol/L/h.

Because of the methodological and unit differences between the two laboratories, enzyme activity values were evaluated separately. For the Vienna cohort, values below 1.2 μmol/L/h were considered reduced. For the Hamburg group, enzyme activity below 140 pmol/spot/21 h was considered low, based on manufacturer guidance. Patients with low enzyme activity according to their respective laboratory criteria were referred for confirmatory genetic testing. GLA gene sequencing was performed using the MiSeq next-generation sequencing platform (Illumina, San Diego, CA, USA) and interpreted according to ACMG guidelines.²³

Statistical Analysis

Data analyses were performed using SPSS version 20.0 (IBM Corp., Armonk, NY, USA). Continuous variables were summarized as mean±standard deviation (SD), while categorical variables were reported as frequencies and percentages. Prevalence of Fabry disease was calculated as a proportion of confirmed cases among the total screened population. Comparisons with other regional prevalence data were made using the Chi-square test, with a p-value <0.05 considered statistically significant.

RESULTS

A total of 613 patients undergoing HD were screened for FD. Of those patients, more than half of them were male (59.9%, n=368) and the mean age was 54.2±15.4 years. The underlying etiologies of ESRD among the screened patients were as follows: diabetic nephropathy (27.6%), hypertensive nephrosclerosis (21.9%), chronic glomerulonephritis (14.8%), autosomal dominant polycystic kidney disease (7.5%), obstructive nephropathy (4.4%), other causes (6.9%), and unknown etiology (16.9%).

Enzyme activity analysis was performed in two reference laboratories using different methods. The first 98 patients were tested at Universitätsklinikum Hamburg-Eppendorf using a DBS assay, where the mean α-Gal A enzyme activity in male patients was 283.4±124.6 pmol/spot/21 h (range: 93.27-700.29). In female patients, the mean α-Gal A enzyme activity was 3.42±1.96 μmol/L/h (range: 0.92-7.85) in the Vienna cohort, and 356.2±128.5 pmol/spot/21 h (range: 110.34-720.44) in the Hamburg cohort. These values were significantly higher compared to male patients, consistent with the X-linked inheritance of Fabry disease. The remaining 515 patients were tested at ARCHIMED Life Science Laboratory (Vienna) using a plasma-based fluorometric assay, with male patients exhibiting a mean α-Gal A enzyme activity of 2.14±1.58 μmol/L/h. Demographic and clinical characteristics are shown in Table.

Table. Demographic and clinical data of the patients	
	Patients (n=613)
Age, years (mean±SD)	54.2±15.4
Gender, male, n (%)	368 (59.9%)
Enzyme values of male patients	
pmol/spot/21h (n=67)	283±124.6
μmol/L/h (n=298)	2.14±1.58
Patients with low α-Gal A, n (%)*	
Male	31 (5.1%)
Female	27 (4.4%)
Patients with mutation, %	2/613 (0.32%)
*Patients with low α-Gal A underwent genetic testing (n=58), SD: Standard deviation, FD: Fabry disease	

Among the 613 screened patients, 96 individuals (15.7%) were identified as having low α-Gal A enzyme activity based on laboratory-specific thresholds. Of these, 58 patients (31 male, 27 female) underwent molecular genetic testing for GLA gene mutations. The remaining patients either declined further evaluation or were lost to follow-up. Enzyme activity in the remaining 517 patients (84.3%) was within the normal reference range.

Two unrelated female patients were found to carry the same novel missense variant, c.116C>A (p.T39K), located in exon two of the GLA gene. The diagnostic flowchart is presented in Figure. This variant was not previously reported in genomAD or ClinVar and no other known pathogenic variants were detected.^{24,25} In silico predictive tools indicated that the p.T39K variant is likely to be functionally damaging. The PolyPhen-2

score was 0.98, classifying the mutation as “probably damaging,” while the SIFT score was 0.01, suggesting the amino acid substitution is “deleterious” to protein function.²⁶ Based on enzymatic deficiency, clinical phenotype, and bioinformatic evidence, the c.116C>A (p.T39K) variant was interpreted as likely pathogenic, supporting the diagnosis of FD in these two patients.

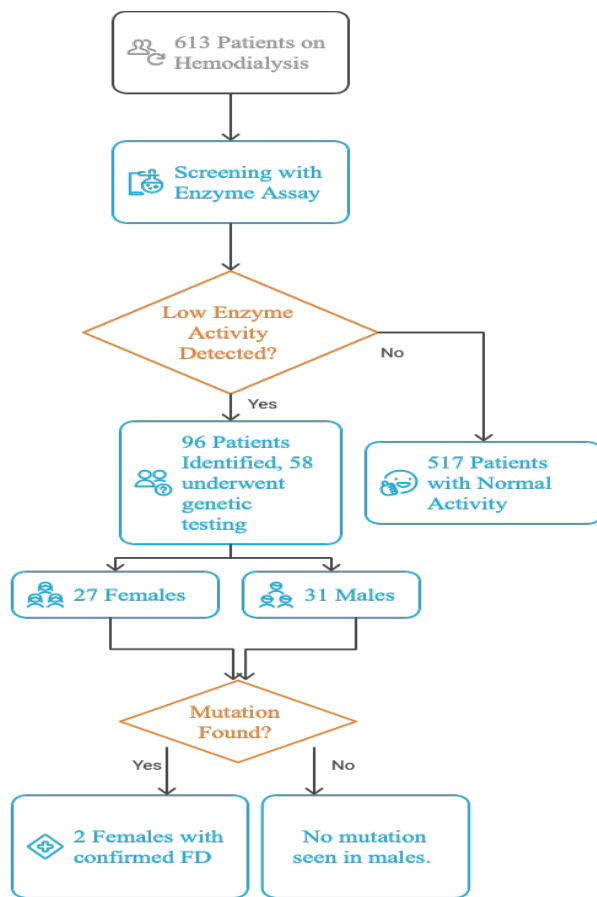


Figure. Schematic flow chart to diagnose FD

Schematic diagnostic flow chart used to identify FD in the study population. Patients were stratified based on laboratory-specific thresholds, and those with low α -Gal A activity were referred for genetic confirmation by GLA sequencing.

FD: Fabry disease

Clinical Features of Mutation-Positive Patients

Case 1: A 45-year-old female patient with ESRD of unknown etiology had been undergoing hemodialysis for five years. She reported intermittent numbness in her hands and feet, as well as mild exercise intolerance. Physical examination was unremarkable with no dermatological abnormalities. Transthoracic echocardiography revealed mild left ventricular hypertrophy (LVH). Genetic analysis identified the c.116C>A (p.T39K) mutation in the GLA gene. In the context of genetic findings, neuropathic symptoms, and cardiac involvement, the patient was classified as having a non-classical Fabry disease phenotype. Family members declined genetic screening.

Case 2: A 38-year-old female patient, previously diagnosed with hypertension and proteinuria, had been undergoing hemodialysis for two years due to ESRD. She reported intermittent numbness in her extremities and occasional dizziness. She also had acroparasthesia. Angiokeratomas

or corneal abnormalities were not observed on physical examination. Transthoracic echocardiography revealed mild LVH. Genetic analysis confirmed the presence of the c.116C>A (p.T39K) variant in the GLA gene. While the patient did not exhibit classical features of FD, the combination of neuropathic symptoms and cardiac involvement was consistent with a possible non classical phenotype. Her family declined genetic testing. Discussion

This study represents the first systematic screening of FD among adult HD patients in Eastern and Southeastern Türkiye, regions with high rates of consanguineous marriage. Among 613 screened patients, two unrelated females were found to carry a previously unreported GLA gene variant, c.116C>A (p.T39K), resulting in an FD prevalence of 0.32%. While this figure aligns with the upper range of prevalence estimates from other high-risk populations, the identification of a novel variant in two patients from a geographically and demographically specific region warrants careful interpretation.

Although Fabry disease is an X-linked disorder and male patients are generally more severely affected, female heterozygotes may also present with significant clinical manifestations.³⁵ In our cohort, mean enzyme activity values in females were higher than those in males, as expected due to random X-chromosome inactivation. However, the wide range observed in female enzyme activities highlights the heterogeneity of biochemical presentation in this group, supporting the need for genetic confirmation in female patients.

The methodological differences between mass spectrometry (DBS samples, Hamburg cohort) and fluorometric assays (plasma samples, Vienna cohort) may have influenced the distribution of enzyme activity values. Previous studies have reported variability in inter-assay correlation, with mass spectrometry providing higher sensitivity for low enzyme activity detection.^{28,29} In our study, enzyme activity thresholds were defined according to laboratory-specific cutoffs, which may limit direct comparison between cohorts. Nonetheless, both methods consistently identified individuals with reduced enzyme activity, and subsequent genetic testing confirmed the absence of false-positive results. This highlights that while methodological differences exist, both testing strategies are reliable for screening dialysis populations, provided that laboratory-specific reference ranges are applied.

Previous studies conducted in patients receiving renal replacement therapy (RRT) have reported the prevalence of Fabry disease to be generally between 0.2% and 0.3%.^{18,27-31} A large screening study from Türkiye reported a prevalence of 0.12% in HD patient group.¹⁸ Consistent with these findings, the prevalence in our study was 0.32%, further supporting the reliability of our screening results. Our findings suggest that regional genetic characteristics-particularly high rates of consanguinity-may play a critical role in the expression and aggregation of rare or novel GLA variants. This is consistent with previous studies that describe the enrichment of private mutations and rare alleles in consanguineous populations due to founder effects and increased homozygosity.^{32,33}

The c.116C>A (p.T39K) variant identified in both patients had not been previously reported in genomic databases such as gnomAD or ClinVar.^{24,25} In silico analyses strongly supported its pathogenicity; PolyPhen-2 classified the substitution as “probably damaging” (score: 0.98), while SIFT predicted it to be “deleterious” (score: 0.01).²⁶ These bioinformatic predictions, in combination with reduced α -Gal A activity and consistent clinical findings support the interpretation of this variant as likely pathogenic. However, the functional impact of the p.T39K variant has not yet been validated through in vitro studies. This step is particularly important, as demonstrated by other variants such as p.D313Y, which was initially classified as pathogenic but later reinterpreted as benign following functional and clinical reassessment.³⁴

Both patients exhibited features suggestive of non-classical or late-onset FD, lacking hallmark signs such as angiokeratomas or corneal opacities, yet presenting with acroparesthesia and cardiac involvement. Such presentations are well documented in female heterozygotes, who may have residual enzyme activity and more attenuated or organ-specific phenotypes.³⁵ These cases highlight the need for genetic confirmation in females, even when enzymatic activity is only mildly reduced, as X-linked inheritance complicates the interpretation of biochemical results.

The study has several important clinical implications. First, these findings support the incorporation of FD screening into routine nephrology protocols in endemic or genetically vulnerable regions. Second, the discovery of a region-specific GLA variant emphasizes the need for geographically tailored screening programs and the development of national mutation registries. Third, early identification of FD provides an opportunity for timely initiation of disease-specific therapies, including enzyme replacement therapy (ERT) or pharmacologic chaperones in amenable genotypes.^{36,37}

Limitations

This study has some limitations. First, no functional validation studies were conducted to assess the impact of the p.T39K variant on enzyme function, folding, or stability. Such analyses are critical for confirming pathogenicity, particularly for novel variants. Second, segregation analysis could not be performed, as family members declined genetic testing. This limited our ability to establish inheritance patterns and strengthen genotype-phenotype correlations. Third, although female patients were included in the enzyme-based screening, α -Gal A activity may remain within normal limits in heterozygous females due to X-chromosome inactivation. As a result, some cases may have been missed. Incorporating direct genetic testing for all female patients could have improved diagnostic accuracy in this group.

CONCLUSION

As a result, this study identified a novel likely pathogenic GLA variant (p.T39K) in two unrelated female HD patients from a high-consanguinity region in Türkiye. These findings highlight the importance of targeted screening in genetically vulnerable populations and suggest that region-specific GLA mutations may contribute to underrecognized cases of FD in nephrology practice. Further functional and family-based

studies are warranted to validate the pathogenicity of the p.T39K variant and to assess its frequency in broader regional cohorts.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Kahramanmaraş Sütcü İmam University Faculty of Medicine Scientific Researches Ethics Committee (Date: 27.07.2015, Decision No: 13).

Informed Consent

All patients signed and free and informed consent form.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Availability of Data and Materials

The data that support the findings of this study are available from the corresponding author, upon reasonable request.

Acknowledgements

We acknowledge the laboratory and clinical staff of the participating centers for their valuable contributions.

REFERENCES

- Desnick R. α -Galactosidase A deficiency: Fabry disease. The metabolic and molecular bases of inherited disease. 1995;2741-84.
- Mauhin W, Lidove O, Masat E, et al. Innate and adaptive immune response in Fabry disease. *JIMD Rep*. 2015;22:1-10. doi:10.1007/8904_2014_371
- Biancini GB, Vanzin CS, Rodrigues DB, et al. Globotriaosylceramide is correlated with oxidative stress and inflammation in Fabry patients treated with enzyme replacement therapy. *Biochim Biophys Acta*. 2012; 1822(2):226-232. doi:10.1016/j.bbdis.2011.11.001
- Arends M, Wanner C, Hughes D, et al. Characterization of classical and nonclassical Fabry disease: a multicenter study. *J Am Soc Nephrol*. 2017;28(5):1631-1641. doi:10.1681/ASN.2016090964
- Kampmann C, Baehner F, Whybra C, et al. Cardiac manifestations of Anderson-Fabry disease in heterozygous females. *J Am Coll Cardiol*. 2002;40(9):1668-1674. doi:10.1016/s0735-1097(02)02380-x
- Shah JS, Elliott PM. Fabry disease and the heart: an overview of the natural history and the effect of enzyme replacement therapy. *Acta Paediatr Suppl*. 2005;94(447):11-10. doi:10.1111/j.1651-2227.2005.tb02103.x
- MacDermot KD, Holmes A, Miners AH. Anderson-Fabry disease: clinical manifestations and impact of disease in a cohort of 98 hemizygous males. *J Med Genet*. 2001;38(11):750-760. doi:10.1136/jmg.38.11.750
- Ortiz A, Oliveira JP, Waldek S, et al. Nephropathy in males and females with Fabry disease: cross-sectional description of patients before treatment with enzyme replacement therapy. *Nephrol Dial Transplant*. 2008;23(5):1600-1607. doi:10.1093/ndt/gfm848

9. Mehta A, Ricci R, Widmer U, et al. Fabry disease defined: baseline clinical manifestations of 366 patients in the Fabry outcome survey. *Eur J Clin Invest.* 2004;34(3):236-242. doi:10.1111/j.1365-2362.2004.01309.x
10. Alroy J, Sabnis S, Kopp JB. Renal pathology in Fabry disease. *J Am Soc Nephrol.* 2002;13(Suppl 2):S134-S138.
11. The Human Gene Mutation Database [Available from:.
12. Germain DP. Fabry disease. *Orphanet J Rare Dis.* 2010;5:30. doi:10.1186/1750-1172-5-30
13. Spada M, Pagliardini S, Yasuda M, et al. High incidence of later-onset fabry disease revealed by newborn screening. *Am J Hum Genet.* 2006;79(1):31-40. doi:10.1086/504601
14. Swinford NA, Prall SP, Gopalan S, et al. Increased homozygosity due to endogamy results in fitness consequences in a human population. *Proc Natl Acad Sci U S A.* 2023;120(43):e2309552120. doi:10.1073/pnas.2309552120
15. Heyne HO, Karjalainen J, Karczewski KJ, et al. Mono- and biallelic variant effects on disease at biobank scale. *Nature.* 2023;613(7944):519-525. doi:10.1038/s41586-022-05420-7
16. Akbayram S, Sari N, Akgün C, et al. The frequency of consanguineous marriage in eastern Turkey. *Genet Couns.* 2009;20(3):207-214.
17. Tunçbilek E, Koc I. Consanguineous marriage in Turkey and its impact on fertility and mortality. *Ann Hum Genet.* 1994;58(4):321-329. doi:10.1111/j.1469-1809.1994.tb00729.x
18. Yalın SF, Eren N, Sinangil A, et al. Fabry disease prevalence in renal replacement therapy in Turkey. *Nephron.* 2019;142(1):26-33. doi:10.1159/000496620
19. Okur I, Ezgu F, Biberoglu G, et al. Screening for Fabry disease in patients undergoing dialysis for chronic renal failure in Turkey: identification of new case with novel mutation. *Gene.* 2013;527(1):42-47. doi:10.1016/j.gene.2013.05.050
20. Kalkan Uçar S, Sozmen E, Duman S, Başçı A, Çoker M. Alpha-galactosidase A activity levels in Turkish male hemodialysis patients. *Ther Apher Dial.* 2012;16(6):560-565. doi:10.1111/j.1744-9987.2012.01092.x
21. Serin EN, Ataş Ü, Çetinkaya R, Sarı F. The Frequency of Fabry disease in hemodialysis patients in the Western Mediterranean region of Türkiye. *Akdeniz Tıp Dergisi.* 2024;10(1):92-99. doi:10.53394/akd.1136791
22. Chamoles NA, Blanco M, Gaggioli D. Fabry disease: enzymatic diagnosis in dried blood spots on filter paper. *Clin Chim Acta.* 2001;308(1-2):195-196. doi:10.1016/s0009-8981(01)00478-8
23. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for molecular pathology. *Genet Med.* 2015;17(5):405-424. doi:10.1038/gim.2015.30
24. Genome Aggregation Database [Available from:.
25. ClinVar [Available from:.
26. Flanagan SE, Patch AM, Ellard S. Using SIFT and PolyPhen to predict loss-of-function and gain-of-function mutations. *Genet Test Mol Biomarkers.* 2010;14(4):533-537. doi:10.1089/gtmb.2010.0036
27. Kotanko P, Kramar R, Devrnja D, et al. Results of a nationwide screening for Anderson-Fabry disease among dialysis patients. *J Am Soc Nephrol.* 2004;15(5):1323-1329. doi:10.1097/01.asn.0000124671.61963.1e
28. Merta M, Reiterova J, Ledvinova J, et al. A nationwide blood spot screening study for Fabry disease in the Czech Republic haemodialysis patient population. *Nephrol Dial Transplant.* 2007;22(1):179-186. doi:10.1093/ndt/gfl528
29. Doi K, Noiri E, Ishizu T, et al. High-throughput screening identified disease-causing mutants and functional variants of α -galactosidase A gene in Japanese male hemodialysis patients. *J Hum Genet.* 2012;57(9):575-579. doi:10.1038/jhg.2012.68
30. Maruyama H, Takata T, Tsubata Y, et al. Screening of male dialysis patients for Fabry disease by plasma globotriaosylsphingosine. *Clin J Am Soc Nephrol.* 2013;8(4):629-636. doi:10.2215/CJN.08780812
31. Herrera J, Miranda CS. Prevalence of Fabry's disease within hemodialysis patients in Spain. *Clin Nephrol.* 2014;81(2):112-120. doi:10.5414/CN108053
32. Szpiech ZA, Xu J, Pemberton TJ, et al. Long runs of homozygosity are enriched for deleterious variation. *Am J Hum Genet.* 2013;93(1):90-102. doi:10.1016/j.ajhg.2013.05.003
33. Xiao Q, Lauschke VM. The prevalence, genetic complexity and population-specific founder effects of human autosomal recessive disorders. *NPJ Genom Med.* 2021;6(1):41. doi:10.1038/s41525-021-00203-x
34. Tomek A, Petra R, Paulasová Schwabová J, et al. Nationwide screening for Fabry disease in unselected stroke patients. *PLoS One.* 2021;16(12):e0260601. doi:10.1371/journal.pone.0260601
35. Izhar R, Borriello M, La Russa A, et al. Fabry disease in women: genetic basis, available biomarkers, and clinical manifestations. *Genes (Basel).* 2023;15(1):37. doi:10.3390/genes15010037
36. Germain DP, Hughes DA, Nicholls K, et al. Treatment of Fabry's disease with the pharmacologic chaperone migalastat. *N Engl J Med.* 2016;375(6):545-555. doi:10.1056/NEJMoa1501098
37. Wanner C, Arad M, Baron R, et al. European expert consensus statement on therapeutic goals in Fabry disease. *Mol Genet Metab.* 2018;124(3):189-203. doi:10.1016/j.ymgme.2018.06.004

The effectiveness of low-dose intravenous ketamine on postoperative pain management in arthroscopic knee surgery under spinal anesthesia

 Dilşad Özkavak¹,  Ayfer Yılmaz²,  Yeliz Kılıç³,  Sadık Özkavak¹,  Nermin Göğüş²

¹Department of Anesthesiology and Reanimation, Eskişehir City Hospital, Eskişehir, Türkiye

²Department of Anesthesiology and Reanimation, Ankara Bilkent City Hospital, Ankara, Türkiye

³Department of Anesthesiology and Reanimation, Faculty of Medicine, Osmangazi University, Eskişehir, Türkiye

Cite this article as: Özkavak D, Yılmaz A, Kılıç Y, Özkavak S, Göğüş N. The effectiveness of low-dose intravenous ketamine on postoperative pain management in arthroscopic knee surgery under spinal anesthesia. *J Health Sci Med.* 2025;8(6):1060-1064.

Received: 20.06.2025

Accepted: 09.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to determine the efficacy of intraoperative low-dose intravenous ketamine on postoperative analgesia in patients who were underwent arthroscopic knee surgery under spinal anesthesia.

Methods: Forty patients were equally classified as control and ketamine groups. In the ketamine group, 0.15 mg/kg intravenous ketamine was administered between 35-45. minutes of the surgery, whereas the patients in the control group received no medication. Postoperative pain was assessed using a visual analogue score. An analgesic, intravenous tramadol, was administered if the patients showed a value of at least 5 on the VAS. The need of first analgesic drug, the time to first analgesic drug administration, and additional (second and third) analgesic requirement during the postoperative period were compared between the groups.

Results: All patients in the control group and 8 (40%) patients in the ketamine group required first analgesic drug ($p<0.001$) postoperatively. The mean time to first analgesic drug administration was longer in the ketamine group (8.5 hours) than in the control group (4.2 hours) ($p=0.008$). Ten patients in the control group needed second analgesic drug whereas only one patient needed the second analgesic in the ketamine group. Third analgesic drug was required in three patients of the control group while none of the patients in the ketamine group needed third analgesic. For patient satisfaction, 11 (55%) patients marked “good” and 9 (45%) marked “very good” in the control group whereas all patients in the ketamine group marked “very good” ($p<0.001$).

Conclusion: Intraoperative subanesthetic dose IV ketamine prolongs the duration of postoperative analgesia and reduces the need for analgesics in elective arthroscopic knee surgery, with an acceptable side effect rates.

Keywords: Arthroscopic knee surgery, ketamine, postoperative analgesia, spinal anesthesia

INTRODUCTION

Spinal block is a frequently used anesthesia method, especially in lower abdominal and extremity surgeries. It is easy to apply, has a rapid onset of action, and provides adequate analgesia and muscle relaxation. Inability to extend the duration of anesthesia in prolonged surgeries due to inadequate block and the need of other methods for pain relief in the postoperative period are its main disadvantages.^{1,2} Postoperative pain is an acute clinical phenomenon that begins with surgical trauma and gradually decreases with tissue healing. It can be influenced with various factors such as the patient's physiological structure and psychological status, the type and duration of surgery, anesthetic drugs and techniques, and the quality of postoperative care. Effective pain management should provide decreased analgesic use in the postoperative period, shorter hospital stay, reduced health costs, and increased patient satisfaction.³

Although many treatment modalities have been recommended for postoperative pain management to date, opioids remain to be among the most commonly used agents in routine practice. However, opioids have numerous adverse effects such as respiratory depression, addiction, drug adaptation, nausea, and vomiting.⁴ On the other hand, ketamine, a noncompetitive N-Methyl-D-Aspartate (NMDA) receptor antagonist, is still used in subanesthetic doses for intra- and postoperative pain control, without major side effects.⁴⁻⁷

In this study, it was aimed to determine the effects of subanesthetic or low-dose intravenous (IV) ketamine administered close to the end of the surgical procedure on hemodynamic parameters, sensory and motor block levels, and postoperative analgesia in patients who were underwent arthroscopic knee surgery under spinal anesthesia.

Corresponding Author: Dilşad Özkavak, dilsadsaracozkavak@hotmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

METHODS

After obtaining approval from Ankara Numune Training and Research Hospital Ethics Committee (Date: 04.04.2007, Decision No: 1), adult patients who were scheduled for elective arthroscopic surgery in Ankara Numune Training and Research Hospital between May 2007 and January 2008 were included in this prospective study. All patients were informed about the operation and anesthesia technique preoperatively. All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Forty patients were randomly and equally divided into two groups: control group and ketamine group, in consecutive order (control group, ketamine group, control group, ...). The patient was blinded about his/her group. The anesthesiologist was also blinded to randomization and performed postoperative tests and assessments without knowledge of the patient group. Between 35-45. minutes of the operation (near the end of the surgery), low (subanesthetic) dose ketamine (0.15 mg/kg bolus IV) was administered in the ketamine group by another anesthesiologist who was fully blinded to study. The patients in the control group received no medication.

The patients less than 18 years old, who had ≥ 3 ASA physical status, central/peripheral neurological disease, drug use that may affect the neurological response, allergy to the ketamine and other anesthetic drugs, and those who had any contraindication to spinal anesthesia were excluded from the study.

Anesthesia Management

Standard monitoring included electrocardiogram (ECG), heart rate (HR), mean arterial pressure (MAP), and peripheral oxygen saturation (SpO_2) values. A 0.9% NaCl infusion (at the rate of 7 ml/kg/hour) was started through the IV access (via 20-gauge venous catheter) at the back of the hand. All patients were premedicated with IV midazolam (0.02 mg/kg).

Spinal anesthesia was performed in the sitting position under appropriate asepsis-antisepsis conditions. A 22-gauge spinal needle was inserted from the midline through the L4-L5 interspinous space, and hyperbaric bupivacaine (0.5%, 2.5 ml) was administered intrathecally after free cerebrospinal fluid flow was observed. Subsequently, the patients were placed in a supine position, and the sensory block levels were checked with pin-prick method. Motor block level was evaluated by Bromage scale (Table 1). When the sensory block level achieved T10, the patients were placed in a semi-sitting position and surgical procedure was initiated. The patients were given 2 L/min oxygen via mask. Standard hemodynamical parameters and sensory block levels were recorded at regular intervals. A 25% decrease in the systolic blood pressure compared with the baseline value was accepted as hypotension, and 10 mg ephedrine was administered intravenously. If the HR was below 50 beats/min, 0.5 mg IV atropine was given. In both groups, hemodynamic parameters, sensory and motor block levels were recorded postoperatively.

Primary and Secondary Endpoints

The effectiveness of subanesthetic IV ketamine administration on postoperative pain status was the primary endpoint of the

present study. Postoperative pain was assessed using a visual analogue score (VAS), scored from 0 (no pain) to 10 (worst pain possible). An opioid analgesic, IV Tramadol (1 mg/kg), was administered if the patients marked a value of 5 or above on the VAS. In this context, the need of "first analgesic drug", the time to "first analgesic drug" administration, additional (second and third) analgesic drug requirements were compared between the groups. The regression time of sensory block to L1, motor block to reach 0, and side effect profiles were the secondary endpoints. Finally, patient satisfaction (poor, good, very good) was evaluated.

Table 1. Motor block levels (Bromage Scale)

0	No motor block
1	Able to flex knee and ankle but unable to lift leg
2	Able to flex ankle but unable to flex hip and knee
3	Complete paralysis

Statistical Analysis

Data analysis was performed using the SPSS (Statistical Package for Social Science) package program. Based on the study by Wu et al.,⁸ the sample size was found at least 36 patients (18 per group) to have a confidence level of 95% that the real value is within $\pm 5\%$ of the measured value. The data were examined for normal distribution using the Shapiro Wilk test. Continuous and categorical variables were shown as mean $\pm$ standard deviation (SD) and number (%), respectively. Whether there was a statistically significant difference between the groups was examined using students T test, Mann-Whitney U test or Chi-square tests. Whether there was a statistically significant difference between repeated hemodynamic measurements within the groups was investigated using Repeated Measurement Analysis of Variance. In cases where the variance analysis result was found to be statistically significant, Bonferroni Corrected multiple comparison test was performed to determine the measurement times that caused the difference. All results were considered statistically significant for $p < 0.05$.

RESULTS

A total of 40 patients were divided as control group ($n=20$) and ketamine group ($n=20$). The groups were statistically similar in terms of age, gender, and operation time. Mean sensory block time to reach T10 and motor block time to reach grade 3 were also similar between the two groups (Table 2).

Table 2. Comparison of basic characteristics and block times between the groups

	Control group	Ketamine group	p
Age (mean, y)	48.9 $\pm$ 9.4	42.1 $\pm$ 14.3	0.398
Gender (F/M)	11/9	8/12	0.342
Operation time (mean, min)	46.1 $\pm$ 10.7	50.4 $\pm$ 8.9	0.102
Sensory block time to reach T10 (mean, min)	6.5 $\pm$ 1.7	7.3 $\pm$ 2.3	0.327
Motor block time to reach grade 3 (mean, min)	5.2 $\pm$ 1.6	6.1 $\pm$ 2.1	0.183

y: year, F: female, M: male, min: minute

In both groups, hemodynamical parameters (HR, MAP, and SpO₂) were recorded several times at regular intervals during the perioperative period. There were no statistically differences in all hemodynamical parameters between the groups ($p>0.05$) (Figure).

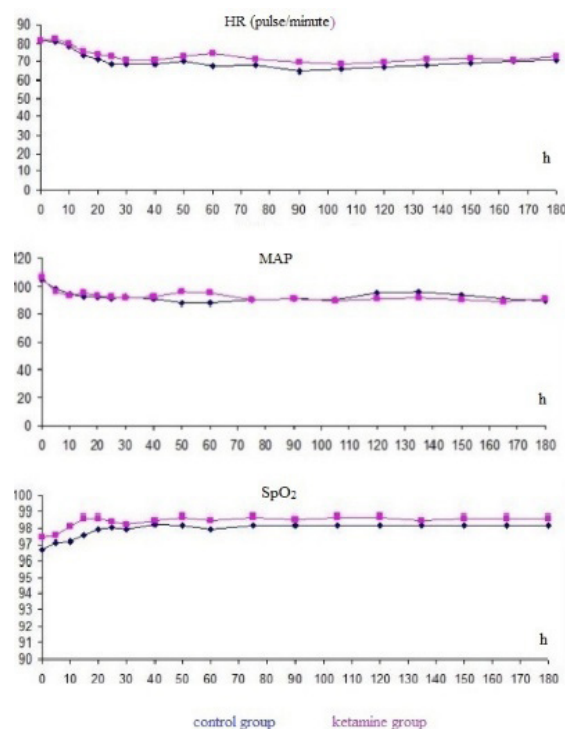


Figure. The courses of HR, MAP, and SpO₂ during the anesthesia period

Postoperatively, the pain levels were assessed with VAS, of which scores of 5 and above was accepted as 'the patient needs analgesic'. In this context, all (100%) patients in the control group and 8 (40%) patients in the ketamine group required "first analgesic drug" ($p<0.001$). The mean time to "first analgesic drug" administration was longer in the ketamine group (8.5 hours) than in the control group (4.2 hours) ($p=0.008$). Ten patients in the control group needed an additional (second) analgesic drug whereas only one patient needed the second analgesic. Thereafter, an additional (third) analgesic drug was required in three patients of the control group while none of the patients in the ketamine group needed third analgesic.

Postoperatively, the regression time of sensory block to L1 was significantly higher in the ketamine group (85.5 ± 24.3 minutes) compared to the control group (40.2 ± 23.7 minutes) ($p<0.001$). There was no statistically significant difference between the groups in terms of the time for postoperative motor block to reach 0 according to the Bromage Scale (111.7 ± 42 minutes in ketamine group vs 109.5 ± 35.7 minutes in control group, $p=0.718$).

In terms of postoperative satisfaction, 11 (55%) patients marked "good" and 9 (45%) marked "very good" in the control group. On the other hand, all patients in the ketamine group marked "very good" ($p<0.001$). No side effects were observed in both groups during the intraoperative period. In the ketamine group, no hallucinations were encountered whereas short-term mild amnesia was observed in all patients.

Nausea and vomiting were observed only in one patient in the ketamine group.

DISCUSSION

This study demonstrated the analgesic efficacy of intraoperative low-dose IV ketamine in arthroscopic surgery under spinal anesthesia. In the postoperative period, all patients in the control group and 8 out of 20 patients in the ketamine group required first analgesic drug. In addition, the time to first analgesic request was 4.2 hours in the control group whereas the time to first analgesic was found 8.5 in the 8 patients of ketamine group, indicating a statistically significance. Moreover, 10 patients in the control group required second analgesic drug while only one patient in the ketamine group was given second analgesic.

As known, ketamine acts multiple sites such as opioid, monoaminergic and muscarinic receptors, induces central summation in second-order pain neurons, thereby provides analgesia by preserving airway patency, ventilation, and cardiovascular stability at subanesthetic doses.^{9,10} Additionally, ketamine reduces the incidence of hypotension induced by spinal anesthesia by increasing heart rate and arterial blood pressure through sympathetic nervous system activation.⁹ In our study, there were no statistical differences in hemodynamic data between the groups. In the literature, there are several studies with conflicting results on the role of ketamine in postoperative pain management when used in low or subanesthetic doses (0.1-0.5 mg/kg) with local anesthetics, opioids or other analgesics.¹²⁻¹⁴ There is a similar study in the literature investigated the effect of low-dose IV ketamine on postoperative pain in knee arthroscopy. In this study, the authors divided patients into three groups; group I (saline control); group II (ketamine 0.15 mg/kg i.v.); and group III (ketamine 0.15 mg/kg + midazolam 0.01 mg/kg i.v.). In conclusion, the authors found that ketamine improved the postoperative pain patient satisfaction.¹⁵

In fact, the studies investigated the role of ketamine on postoperative pain management have focused on two important issues; the administration time (pre/intra/postoperative) and the dose.¹⁶ For example, Wu et al.,⁸ showed that preincisional analgesia with morphine, bupivacaine, and ketamine was better than postincisional application. In parallel, Samuel et al.⁴ reported that low-dose IV ketamine before skin incision was extended postoperative first analgesia request time and decrease total analgesia consumption in 24 hours. In another study conducted on the patients who underwent total hip arthroplasty, the authors demonstrated the analgesic efficacy of both low-dose preincisional and a 24-hour infusion IV ketamine.¹⁷ On the other hand, Najam and Jafri⁷ compared the postoperative analgesic effects of intradermal lidocaine with IV ketamine in patients who underwent elective inguinal hernia repair under spinal anesthesia. The authors gave the patients 50 mg bolus ketamine after the completion of the surgery. Based on their results, administering bolus ketamine significantly decreased overall pain scores and reduce the incidence of chronic postsurgical pain, compared to lidocaine infiltration. In another study, intraoperative low-dose ketamine (0.15 mg/kg)

was determined to reduce the need of postoperative morphine and accelerate mobilization after 24 hours; however, the time of ketamine administration (before incision or at the end of surgery) did not affect the outcomes.¹⁸ As a result, ketamine was given until the end of surgery in most studies, although the timing of administration varied. In parallel, ketamine was administered near the end of surgery in our study.

Another important issue in the use of ketamine is the dosage. At high doses, ketamine provides both anesthesia and analgesia, and is often used in diagnostic and surgical procedures that do not require skeletal muscle relaxation. Ketamine is used as an adjunct to other anesthetic drugs because of its ability to maintain heart rate and cardiac output, to increase blood pressure via sympathetic stimulation, to provide bronchodilation, and to preserve protective pharyngeal and laryngeal reflexes without depressing respiration.¹⁹ Its safety profile also makes it suitable for anesthesia and sedation in areas with limited medical equipment. However, the fact that ketamine may cause a range of psychosomatic symptoms, such as delusions, hallucinations, delirium, and confusion, greatly limits its widely use.²⁰ On the other hand, ketamine acts as an effective analgesic and causes fewer psychotomimetic side effects at low or subanesthetic doses.²¹ In a randomized double-blinded study investigated the effect of low-dose intraoperative ketamine on postoperative pain after total knee replacement surgery, the authors reported decreased pain scores, prolonged time to first analgesic request, and reduction in morphine consumption, without affecting the incidence of side effects.²² Similarly, low-dose ketamine was used in our study cohort, and no hallucinations or other serious psychotomimetic symptoms were observed during the postoperative period. However, all patients in the ketamine group experienced amnesia which was mild, transient, and clinically insignificant. Indeed, the development of psychotomimetic symptoms is dose-dependent in ketamine use, and mild amnesia is a common and well-known side effect in its subanesthetic doses.²³ The factors limiting the use of high doses of ketamine also include postoperative nausea, vomiting, and visual disturbances such as blurred vision and nystagmus.^{24,25} In the current study, no patient developed any visual problem, and only one patient in the ketamine group experienced mild vomiting in the recovery room.

As the secondary outcomes of the present study, the regression time of sensory block was prolonged in ketamine group whereas the duration of motor block and the onset times of motor and sensory blocks were found to be similar between the groups. These results were consistent with a recent meta analysis regarding the use of ketamine as an adjuvant to regional anesthesia.²⁶ This large-scale study clearly showed that ketamine provides advantages in postoperative pain control through different pain pathways and by prolonging the sensory block time.

Limitations

There were several limitations to our study. First, all blocks were performed by a single anesthesiologist in a single center, which facilitated standardization but limited the generalizability of the results. Second, the relatively small sample of subgroups may be considered another limitation of

this work. The fact that no medication (placebo) was given to the control group and that side effects such as hallucinations, amnesia, nausea, and vomiting were assessed only clinically can also be considered as another limitations of this study. Finally, body-mass index, which is an important factor affecting drug metabolism and may influence the results of the study, was not compared between the groups because of the irregular records. This can be considered as another limitation.

CONCLUSION

The results obtained from the study showed that the administration of subanesthetic dose IV ketamine, as an adjuvant to intrathecal bupivacaine provided longer postoperative analgesia and less need of postoperative analgesic use in elective arthroscopic surgery, with an acceptable side effect profile.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Ankara Numune Training and Research Hospital Ethics Committee (Date: 04.04.2007, Decision No: 1).

Informed Consent

All patients signed and free and informed consent form.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Vadhanan P. Recent updates in spinal anesthesia-a narrative review. *Asian J Anesthesiol.* 2021;59(2):41-50. doi:10.6859/aja.202106_59(2).0001
2. Ledesma I, Stieger A, Luedi MM, Romero CS. Spinal anesthesia in ambulatory patients. *Curr Opin Anaesthesiol.* 2024;37(6):661-665. doi:10.1097/ACO.0000000000001412
3. Dimaculangan D, Chen JF, Borzio RB, Jauregui JJ, Rasquinha VJ, Maheshwari AV. Periarticular injection and continuous femoral nerve block versus continuous femoral nerve block alone on postoperative opioid consumption and pain control following total knee arthroplasty: randomized controlled trial. *J Clin Orthop Trauma.* 2019;10:81-86. doi:10.1016/j.jcot.2017.09.012
4. Samuel H, Aweke S, Tuni J. Effect of low-dose intravenous ketamine on postoperative pain following cesarean section under spinal anesthesia: a prospective cohort study, Ethiopia. *Ann Med Surg (Lond).* 2022;77:103570. doi:10.1016/j.amsu.2022.103570
5. Adhikari P, Subedi A, Sah BP, Pokharel K. Analgesic effects of intravenous ketamine after spinal anaesthesia for non-elective caesarean delivery: a randomised controlled trial. *BMJ Open.* 2021;11(6):e044168. doi:10.1136/bmjopen-2020-044168

6. Gemechu AD, Gebremedhin TD, Andebiku AA, Solomon F, Sorsa A. The effect of ketamine versus tramadol on prophylactic post-spinal shivering in those patients undergoing orthopedic surgery: a prospective cohort study design, 2020. *BMC Anesthesiol.* 2022;22(1):361. doi:10.1186/s12871-022-01906-z
7. Fahad Najam M, Jafri N. A Comparison between the effectiveness of ketamine bolus and intradermal lidocaine in reducing acute postoperative pain. *Cureus.* 2022;14(7):e26563. doi:10.7759/cureus.26563
8. Wu CT, Yeh CC, Yu JC, et al. Pre-incisional epidural ketamine, morphine and bupivacaine combined with general anesthesia provides pre-emptive analgesia for upper abdominal surgery. *Acta Anesthesiol Scand.* 2000;44:63-68. doi:10.1034/j.1399-6576.2000.440112.x
9. Tawfic QA. A review of the use of ketamine in pain management. *J Opioid Manag.* 2013;9:379-388. doi:10.5055/jom.2013.0180
10. Mahdi AH, Kahloul M, Mohammed MJ, Mohammed AK. Effects of ketamine and tramadol as adjuvants to bupivacaine in spinal anesthesia for unilateral open ovarian cystectomy: a randomized controlled trial. *Cureus.* 2024;16(2):e54776. doi:10.7759/cureus.54776
11. Aboelsuod MAA, Elnaggar AMA, Alwafa TAAA, Ahmed MMH, Elbeltagy ASA, Elbarbary MIA. Effect of intravenous ketamine infusion on hemodynamics of patients undergoing cesarean delivery after spinal anaesthesia: a randomized, double-blind, controlled trial. *Turk J Anaesthesiol Reanim.* 2023;51(5):420-426. doi:10.4274/TJAR.2023.231231
12. Guo J, Zhao F, Bian J, Hu Y, Tan J. Low-dose ketamine versus morphine in the treatment of acute pain in the emergency department: a meta-analysis of 15 randomized controlled trials. *Am J Emerg Med.* 2024;76:140-149. doi:10.1016/j.ajem.2023.11.056
13. Xiang J, Cao C, Chen J, Kong F, Nian S, Li Z, Li N. Efficacy and safety of ketamine as an adjuvant to regional anesthesia: a systematic review and meta-analysis of randomized controlled trials. *J Clin Anesth.* 2024;94:111415. doi:10.1016/j.jclinane.2024.111415
14. Fuller RG, Kikla EM, Fawcett APW, et al. Low-dose ketamine for acute pain: a narrative review. *Am J Emerg Med.* 2024;86:41-55. doi:10.1016/j.ajem.2024.09.033
15. Cagla Ozbakis Akkurt B, Inanoglu K, Kalaci A, Turhanoglu S, Asfuroglu Z, Tumkaya F. Effects of intravenous small dose ketamine and midazolam on postoperative pain following knee arthroscopy. *Pain Pract.* 2009;9:289-295. doi:10.1111/j.1533-2500.2009.00278.x
16. Ma X, Yan J, Jiang H. Application of ketamine in pain management and the underlying mechanism. *Pain Res Manag.* 2023;2023:1928969. doi:10.1155/2023/1928969
17. Remérand F, Le Tendre C, Baud A, et al. The early and delayed analgesic effects of ketamine after total hip arthroplasty: a prospective, randomized, controlled, double-blind study. *Anesth Analg.* 2009;109(6):1963-1971. doi:10.1213/ANE.0b013e3181bdc8a0
18. Menigaux C, Fletcher D, Dupont X, Guignard B, Guirimand F, Chauvin M. The benefits of intraoperative small-dose ketamine on postoperative pain after anterior cruciate ligament repair. *Anesth Analg.* 2000;90(1):129-135. doi:10.1097/00000539-200001000-00029
19. Peltoniemi MA, Hagelberg NM, Olkkola KT, Saari TI. Ketamine: a review of clinical pharmacokinetics and pharmacodynamics in anesthesia and pain therapy. *Clin Pharmacokinet.* 2016;55(9):1059-1077. doi:10.1007/s40262-016-0383-6
20. Morgan CJ, Curran HV; Independent Scientific Committee on Drugs. Ketamine use: a review. *Addiction.* 2012;107(1):27-38. doi:10.1111/j.1360-0443.2011.03576.x
21. Kwok RFK, Lim J, Chan MTV, Gin T, Chiu WKY. Preoperative ketamine improves postoperative analgesia after gynecologic laparoscopic surgery. *Anesth Analg.* 2004;98(4):1044-1049. doi:10.1213/01.ANE.0000105911.66089.59
22. Cengiz P, Gokcinar D, Karabeyoglu I, Topcu H, Cicek GS, Gogus N. Intraoperative low-dose ketamine infusion reduces acute postoperative pain following total knee replacement surgery: a prospective, randomized double-blind placebo-controlled trial. *J Coll Physicians Surg Pak.* 2014;24(5):299-303.
23. Orhurhu VJ, Vashisht R, Claus LE, Cohen SP. Ketamine Toxicity. 2023. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025.
24. Gorlin AW, Rosenfeld DM, Ramakrishna H. Intravenous sub-anesthetic ketamine for perioperative analgesia. *J Anaesthesiol Clin Pharmacol.* 2016;32(2):160-167. doi:10.4103/0970-9185.182085
25. Wang X, Lin C, Lan L, Liu J. Perioperative intravenous S-ketamine for acute postoperative pain in adults: a systematic review and meta-analysis. *J Clin Anesth.* 2021;68:110071. doi:10.1016/j.jclinane.2020.110071
26. Xiang J, Cao C, Chen J, et al. Efficacy and safety of ketamine as an adjuvant to regional anesthesia: a systematic review and meta-analysis of randomized controlled trials. *J Clin Anesth.* 2024;94:111415. doi:10.1016/j.jclinane.2024.111415

Examining jigsaw, metaverse, and classic learning methods' impact on nursing students' self-efficacy in pediatric drug administration: a mixed-methods study

Şerife Tutar, Hande Özgörü, Faruk Durna, Yurdagül Şahin

Department of Pediatric Nursing, Faculty of Health Sciences, Süleyman Demirel University, Isparta, Türkiye

Cite this article as: Tutar Ş, Özgörü H, Durna F, Şahin Y. Examining jigsaw, metaverse, and classic learning methods' impact on nursing students' self-efficacy in pediatric drug administration: a mixed-methods study. *J Health Sci Med.* 2025;8(6):1065-1075.

Received: 14.08.2025

Accepted: 09.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to examine the effects of jigsaw, metaverse, and classic learning methods on nursing students' self-efficacy levels in pediatric drug administration and to explore students' perspectives on these learning methods.

Methods: A mixed method study was conducted with 150 students in Türkiye. The data were collected "Introductory Information Form", "Medication Administration Self-Efficacy Scale in Children for Nursing Students" and "Semi-Structured Interview Form". Our research was performed in two phases. In the first stage, the self-efficacy levels of the students in drug administration in children were determined using a randomized controlled research design with pre-test and post-test administration. In the second stage, face-to-face qualitative interviews were conducted with 10 volunteer students from each learning method (jigsaw, metaverse, classic) about the positive and negative aspects of the learning method.

Results: The majority of the students participating in the study were female (75.3%), with a mean age of 21.98 ± 2.23 years. A significant difference was determined in the post-test MASES-C scores between the groups in which the jigsaw, metaverse, and classic learning methods were applied (71.51 ± 4.99 ; 67.74 ± 6.29 ; 68.66 ± 7.70 , respectively). The positive aspects of the learning methods included interactive course delivery and reduced instructor-induced stress (jigsaw), time-saving and flexibility (metaverse), and direct interaction with the instructor (classic learning). The negative aspects were technical issues and distractions (metaverse) and monotony and one-way communication leading to reduced attention span (classic learning).

Conclusion: Our research reveals that jigsaw, metaverse, and classic learning methods can enhance nursing students' self-efficacy in pediatric drug administration. Among these, jigsaw learning provided the most effective learning environment, yielding the highest self-efficacy scores.

Keywords: Learning methods, nursing education, self-efficacy, pediatric nursing

INTRODUCTION

Patient safety is a global health concern, and medication errors are the most common mistakes among pediatric nurses.¹⁻³ Due to differences in metabolism, absorption, and growth processes, drug selection and dosage must be carefully adjusted in pediatric patients.⁴ Medication errors are more frequent in pediatric patients than in adults.² To prevent these errors, nursing students must acquire sufficient knowledge and skills before graduation. Given that students have a technology-driven lifestyle, educational approaches should integrate innovative technologies rather than relying on traditional methods.^{5,6}

Global advancements in nursing science and health technology are transforming knowledge production and application, increasing the need for adaptable, quality-focused nurses.⁷ In this rapidly changing world, educators should adopt interactive, modern teaching approaches to enhance student development.⁸ Instead of traditional, rote-based learning,

education should encourage active student participation. This has highlighted the jigsaw learning method, which enhances self-confidence, communication, critical thinking, ethical decision and problem-solving skills.^{5,7,9} In this method, students form expert groups on specific topics, then share their expertise with their original groups.⁷ The jigsaw technique fosters idea-sharing and continuous engagement, positively impacting nursing students' academic and problem-solving skills.⁹

The COVID-19 pandemic and the subsequent earthquake disaster in Türkiye highlighted the urgent need for flexible and resilient education systems. These extraordinary circumstances accelerated the transition from traditional to online learning, leading to the growing use of innovative approaches such as the metaverse. The metaverse enables immersive and interactive learning in virtual environments, allowing continuity of education during crises by maintaining interaction and

Corresponding Author: Şerife Tutar, serifeguvn@sdu.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

engagement.¹⁰⁻¹³ Despite its advantages, however, challenges such as limited social interaction, technical problems, and ethical concerns have been reported.^{12,14,15} Previous studies demonstrated that metaverse-based learning could improve outcomes in nursing education. For example, Cho et al.¹⁶ developed a family-centered metaverse simulation that enhanced nursing students' self-efficacy in patient delivery, while Kim et al.¹⁷ emphasized that avatar-based interactions alone may not fully support a sense of spatial presence. These findings suggest that integrating diverse learning strategies may enhance effectiveness. Nevertheless, current evidence on the use of innovative methods such as metaverse and active learning techniques in nursing education remains limited. Especially in pediatric nursing education, the evidence regarding the use of these methods is highly restricted, which not only underscores the importance of further research but also highlights the need for scientific data to guide decisions on appropriate teaching methods, the planning of skill-based training, and the preparation of students for clinical practice.^{5,7,16} Therefore, the aim of this study is to examine the effects of jigsaw, metaverse-based, and traditional learning methods on nursing students' self-efficacy in pediatric drug administration and to explore their perspectives on these learning approaches. The findings are expected to guide the adoption and implementation of innovative teaching techniques in nursing education, particularly within the field of pediatric nursing.

METHODS

Ethical Considerations

Approval was obtained from the Süleyman Demirel University Faculty of Medicine Clinical Researches Ethics Committee to conduct the research (Date: 13.04.2023, Decision No: 87), and institutional permission was also obtained. Written informed consent was obtained from all students who participated in the study. All procedures were carried out in accordance with the ethical rules and the principles.

Design

This research was conducted using a mixed-methods design that combined both quantitative and qualitative approaches. The good reporting of a mixed methods study (GRAMMS) guideline,¹⁸ CONSORT,¹⁹ and the consolidated criteria for reporting qualitative studies (COREQ)²⁰ were followed to ensure methodological rigor. The study was carried out in two phases. In the first phase, a randomized controlled trial with a pre-test/post-test design was conducted to evaluate nursing students' self-efficacy in pediatric drug administration (NCT06836674). In the second phase, purposive sampling was used to ensure representation from each learning group (jigsaw, metaverse, classic), and semi-structured, face-to-face interviews were conducted with 10 volunteer students per group (n=30 in total). The number of participants was determined based on the principle of data saturation rather than statistical power. Previous literature suggests that 10-15 participants are often sufficient to obtain rich and meaningful data in in-depth interviews, and in our study, saturation was reached when no new codes or themes emerged.²¹ The interview guide included open-ended questions that explored students' perceptions of the advantages, challenges, and

overall experiences of the learning methods. Interviews were conducted in a quiet classroom environment, lasted 20-30 minutes, and were audio-recorded with participants' consent. Audio recordings were transcribed verbatim, organized using MAXQDA Analytics Pro 2024, and analyzed thematically following Braun and Clarke's approach.²² Two independent researchers carried out the coding process. To ensure trustworthiness, the findings were discussed with a faculty member who was independent of the study, and member checking was conducted; these strategies helped to verify the accuracy of the results and reduce the risk of researcher bias.

The hypotheses and research questions of the research were as follows:

Hypotheses

- **H0:** There would be no significant difference between the scores of jigsaw, metaverse, and classic learning groups in the pre-test and post-test scores of pediatric drug administration self-efficacy levels.
- **H1:** There would be a significant difference between the pre-test and post-test scores of the self-efficacy levels of pediatric drug administration, jigsaw, metaverse, and classic learning.

Research Question

- **Q1:** What are the positive and negative aspects of jigsaw, metaverse, and classic learning methods applied in pediatric drug administration skills training?
- **Q2:** What are the recommendations of students for pediatric drug administration skills training?

Setting and Sample

The research was conducted between 08 May-09 June 2023 at a university providing nursing education in Türkiye. The population of the study consisted of 187 students who took the Pediatric Health and Diseases Nursing Course of the Department of Nursing at this university. Power analysis was performed to determine the sample number of the study. Power analysis of the research G*power is (Heinrich Heine University, Dusseldorf, Germany). Since the design of the study was prepared as three independent groups and two repeated measurements, "F test" was selected as the test family and "ANOVA repeated measures, within-between interaction" was selected as the analysis. For repeated measurements, the variance explained by the effect was taken as 1 and the effect size was calculated as $f=0.312$. The total sample size was calculated as $n=135$ for three independent groups and two repeated measurements under the conditions of 5% error margin, 0.95 power value, and $r=0.5$ correlation between repeated measurements. However, considering the randomization and repeated study situation, it was decided to complete it with a total of $n=150$ students, 50 for each group. There was no sample loss during the research process.

Inclusion criteria for in the study: Taking the Child Health and Diseases Nursing Course for the first time, and volunteering to participate in research.

Exclusion criteria for from the study: Not attending class during the study data collection dates, not having previously

taken courses/seminars on pediatric drug administration, and failing the course the previous year and taking the course for the second time.

Randomization

In our study, the students who met the sampling inclusion criteria were divided into three groups using the block randomization method, in which the number and probability of individuals in each block were equal (Figure 1). For this purpose, a complete list of eligible students was created, and each student was assigned a unique identification number. Block sizes were determined to ensure equal distribution across the three intervention arms. Using a computer-based randomization program (<https://www.randomizer.org/#randomize>), the independent statistician generated random sequences and allocated the identification numbers to one of the three groups (jigsaw, metaverse, or classic learning). This procedure ensured that each group included an equal number of participants (n=50), maintaining balance and minimizing selection bias. The randomization process was performed entirely by the statistician who conducted the statistical analyses and was not involved in data collection or intervention delivery.

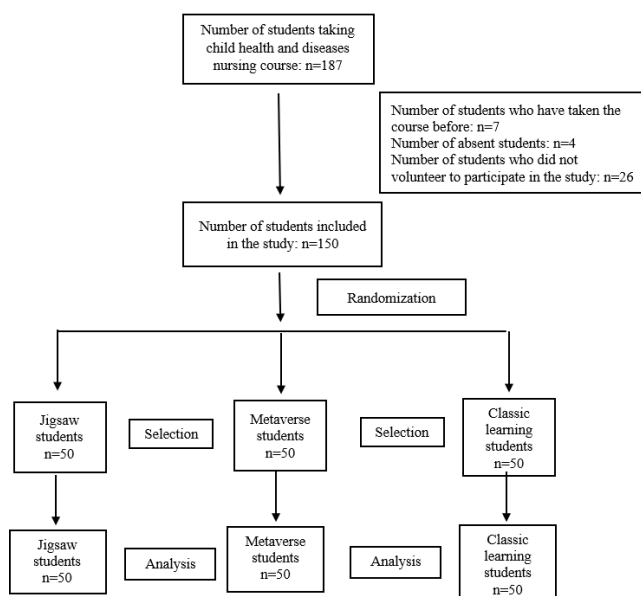


Figure 1. CONSORT flow diagram

Blinding

In the learning methods used in our research, blinding could not be achieved because the students had to actively participate in the learning processes and the instructors were aware of the method they were delivering. However, to minimize bias, the randomization scheme and the entire statistical analysis were conducted by an independent statistician who was blinded to the group assignments. Thus, blinding was applied only at the level of statistical analysis, not during the intervention delivery.

Data Collection Tool

The data of the study were collected using the Introductory Information Form, the Medication Administration Self-Efficacy Scale in Children for Nursing Students (MASES-C),

and the Semi-Structured Interview Form prepared for use in qualitative interviews.

Introductory Information Form: The information form created by the researchers included information such as age, sex, type of high school from which the student graduated, and previous participation in group trainings.

Medication Administration Self-Efficacy Scale in Children for Nursing Students (MASES-C): MASES-C was developed by Bektaş et al.²³ in 2021. The scale consists of 16 items and two sub-dimensions. The drug preparation sub-dimension for children (items for calculating and obtaining drug doses) includes items 1-8, and the drug administration sub-dimension (items for administering the obtained drug) comprises items 9-16. This scale is rated in the Likert type of 5 and is listed as 1=Absolutely not sufficient, 2=Somewhat sufficient, 3=Partially sufficient, 4=Sufficient, 5=Absolutely sufficient. The lowest score from the scale is 16 and the highest score is 80. The scale has no cut-off point, the higher the score, the higher the students' self-efficacy regarding pediatric drug practice. There are no inverted score items on the scale.

Semi-Structured Interview Form: This form, prepared by the researchers, consisted of three open-ended questions ("What are the positive aspects of the learning method you are involved in?", "What are the negative aspects of the learning method you are involved in?", "What are your suggestions and wishes regarding learning methods?") and open-ended sub-questions according to the course of the interview.

Data Collection

In our research, first, the topics and content of the subject of "pediatric drug administration training" were created by the researchers and presented for expert opinion, and the training content was finalized in line with the expert opinions. Afterward, the training content was made ready to be shared with the student according to classic, metaverse-based, and jigsaw learning methods. The training content was prepared under five main headings in total and with five sub-headings under each main heading. The training content was presented on the same day in the same week in accordance with the learning method of each group. Before the training content started to be presented, a face-to-face pre-test was applied to all groups simultaneously on MASES-C.

Jigsaw technique: The Jigsaw I technique was used in this learning method. Because the subject of pediatric drug administration was grouped under five main topics and each main topic under five headings, students formed main learning groups of five people before the application to share the subtopics within the main topic. Each student was assigned a number between 1-5 in the group.

Educational content prepared by the responsible researcher in line with the opinions of experts was distributed to each group and guidance was provided to the students on how to work. Each student was assigned a sub-topic of a main topic. All the main groups first came together in a laboratory environment to learn their subjects. Then, students who took the same sub-subject in the main learning groups gathered and formed jigsaw (expertise) groups. After the work in the jigsaw

group was completed, the students returned to their original learning groups and taught their expertise to the other group mates. After the study of all groups was completed, the post-test MASES-C was applied to the students to determine the knowledge levels of the students.

Metaverse-based education: In our research, some administrations were used in the process of creating the metaverse universe. Students were provided with an interactive and customized learning experience using technologies such as Blender 3D modeling, virtual reality (VR), augmented reality (AR), artificial intelligence (AI), cloud computing, and blockchain (Figure 2). The basic infrastructure of education is built on the platform of Spatiol.io, which enabled students to participate in virtual classrooms and access educational materials using their virtual avatars.²⁴ Through this platform, our students were given the opportunity to learn by doing group work, individual interactions, and watching educational content, and the processes of pediatric drug administration were tried to be made more effective and understandable. This is in line with recent studies showing that metaverse platforms can enhance collaboration and experiential learning in health sciences education.²⁵⁻²⁷

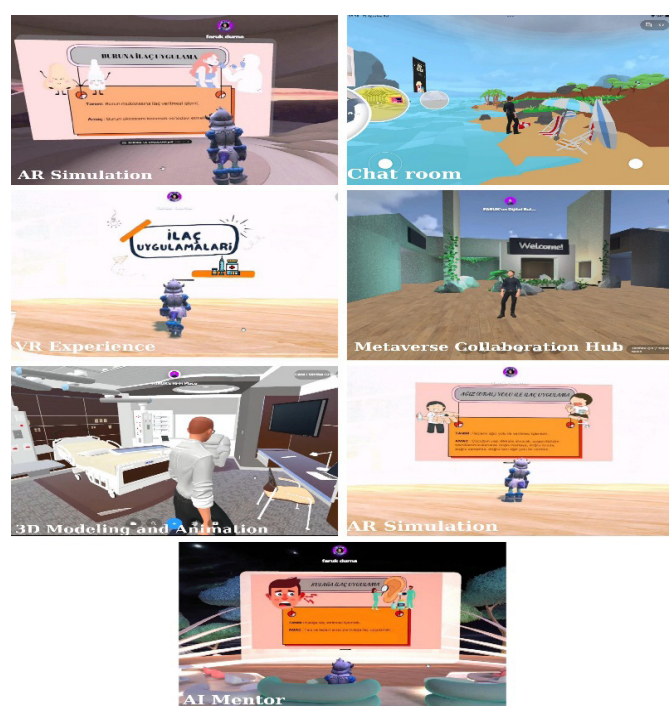


Figure 2. Screenshots of the virtual classroom and chat room in the Metaverse environment

Blender software was preferred for the visualization of the training materials on the platform. With this software, it is aimed to make the learning processes more effective and understandable by supporting the training for nursing students in pediatric drug administration using 3D models. Another important element of our research infrastructure is the integration of VR and AR technologies using Unity. Unity's powerful development tools enabled students to step into a fully virtual world using VR headsets, and AR technology allowed digital content to integrate with the physical world.²⁸ In this way, theoretical knowledge was tried to be transformed into a practical learning experience. Previous

reviews have highlighted that immersive VR/AR simulations improve student engagement, motivation, and clinical skills acquisition in nursing and medical education.^{29,30}

In our research, AI served as a digital mentor guiding students. AI algorithms analyzed students' learning processes, provided content appropriate to their needs, and helped them address their individual deficiencies. In addition, with natural language processing (NLP) technology, intelligent chatbots were developed that instantly respond to students' questions about pediatric drug administration. Cloud computing, on the other hand, enabled all data to be securely stored and managed in a central system. This technology enabled us to rapidly scale and make available student information, learning materials, and virtual environments.³¹ Evidence also suggests that cloud-based infrastructures improve scalability and accessibility in health sciences and biomedical education.³² Blockchain technology was used to increase transparency and comprehensibility in pediatric drug administration. It also created a verifiable system by securely recording student performances and educational records using blockchain.

In terms of pedagogical design, the metaverse intervention was structured with clear educational objectives focusing on safe pediatric drug preparation and administration. Students were able to access the virtual classroom simultaneously, and there was no fixed instructor–student ratio or strict time limit, as the platform allowed self-paced exploration in addition to synchronous group interactions. Each session included virtual demonstrations, case-based scenarios, group discussions, and opportunities for students to practice drug administration steps with immediate feedback from the system. The total intervention period was matched with the other groups, but within the metaverse environment students had the flexibility to revisit materials, repeat activities, and interact with peers and digital mentors as needed. The created course content was made available to the students in the metaverse group on the same day as the other groups.

Classic learning: The content, which was prepared in line with the expert opinions in the classic learning method, was converted into slide presentations and presented to the students by the responsible researcher and transferred to the students in the laboratory environment for a total of 4 lesson hours consisting of 45 minutes of class time each week.

Statistical Analysis

The data obtained in the study were analyzed using the SPSS is IBM SPSS Statistics version 25 (IBM Inc., Armonk, NY, USA). Descriptive statistical methods were used to evaluate the data. The Levene test was used to test the homogeneity of the groups. The conformity of continuous numerical data to normal distribution was tested using Kolmogorov-Smirnov analysis. One-way analysis of variance (ANOVA) was applied for the difference between more than two independent groups in the comparison of quantitative data in normally distributed data, and the paired T test was applied in the comparison of two dependent measurements. In the comparisons of categorical variables according to the study groups, the Monte Carlo corrected Chi-square test was used. In all analyses, p-values of <0.05 were considered statistically

significant for a type-I error rate of 5%. The responses to the open-ended questions were evaluated using a descriptive thematic analysis method. In this process, two researchers first read all the responses several times to gain familiarity with the data and to develop a comprehensive perspective. Subsequently, meaningful text segments were identified and initial codes were generated inductively. Similar codes were then grouped under broader categories, which were refined into themes based on their frequency, representativeness, and relevance to the research questions. To ensure the reliability of the coding process, two researchers conducted coding independently, discussed any discrepancies, and reached consensus, thereby strengthening inter-coder consistency. MAXQDA Analytics Pro 2024 was used to organize and manage the data; however, theme development was carried out in line with methodological principles rather than being limited to the software. In the final stage, the identified themes were supported with representative quotations from participants and continuously compared with the raw data to ensure coherence and accuracy.

RESULTS

Findings on the Sociodemographic Characteristics of the Students

A total of 150 nursing students were included in the study. The majority of the students were female (75.3%) and the overall average age was calculated as 21.98 ± 2.23 years. The vast majority of the students (83.3%) were determined as Anatolian/Science high school graduates and the rest were determined as plain high school graduates. Each research group consisted of 50 students. The proportion of students who received group education in previous courses was found to be 67.3%. There was no significant difference between the mean age, sex, and high schools of the students in the study groups, and the groups showed a homogeneous distribution (Table 1).

Findings on the scores of the students from the self-efficacy scale and scale sub-dimensions of drug administration in children: Pre-test and post-test values were compared

in each study group. For the jigsaw study group, both drug preparation and drug administration sub-dimensions and drug administration self-efficacy scores showed significant differences between the pre-test and post-test ($p < 0.001$). It was observed that the pre-test scores were significantly lower than the post-test scores. Similarly, drug preparation, drug administration, and general self-efficacy scores differed significantly for both pre- and post-test in the metaverse and classic learning groups ($p < 0.001$). Post-test scores were found to be significantly higher in all study groups.

Drug preparation, drug administration, and general self-efficacy pretest values of the drug administration self-efficacy scale did not differ significantly between the study groups ($p > 0.05$). However, the post-test value of the drug preparation sub-dimension did not reach statistical significance among the study groups ($p = 0.074$). The post-test value of the drug application dimension was found to have a higher score in the learning group using the jigsaw method ($p = 0.038$). Thus, MASES-C scores were found to be higher in the students using the jigsaw method (71.51 ± 4.99) and lower in the students using the metaverse method (67.74 ± 6.29) ($p = 0.029$) (Table 2). The values related to the scale scores are presented in Figure 3.

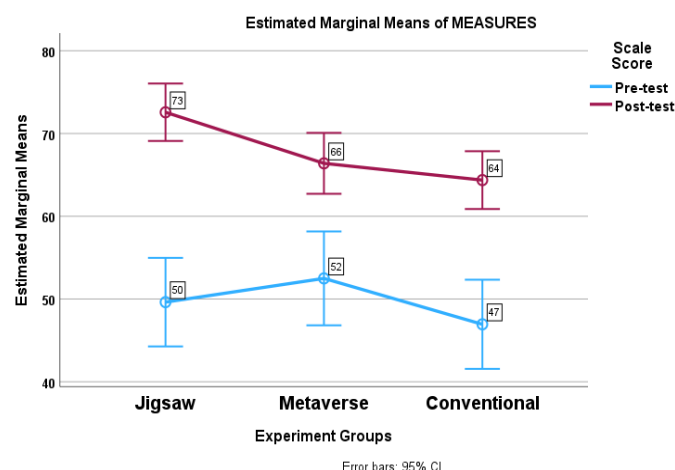


Figure 3. Drug preparation self-efficacy scores pre-test-post-test values according to study groups

Table 1. Distribution of the participants participating in the research according to their demographic characteristics

Properties	Jigsaw learning		Metaverse-based learning		Classic learning		Test value	p
	n	%	n	%	n	%		
Sex								
Female	35	70.0	42	84.0	36	72.0	3.244	0.197**
Male	15	30.0	8	16.0	14	28.0		
Education								
Anatolian/science high school	40	80.0	44	88.0	41	82.0	1.296	0.523**
Plain high school	10	20.0	6	12.0	9	18.0		
Status of receiving group training								
Yes	35	70.0	36	72.0	30	60.0	1.855	0.396**
No	15	30.0	14	28.0	20	20.0		
Age (years)								
(X±SD)	22.18±2.17		21.70±2.01		22.08±2.51		0.637	0.530*
Total	50	100.0	50	100.0	50	100.0		
*Student T tests, **Chi-square analysis, SD: Standard deviation								

*Student T tests, **Chi-square analysis, SD: Standard deviation

Table 2. Comparison of participants' self-efficacy scale and dimensions for medication administration in children among nursing students between and within groups

		Pre-test		Post-test			
Scale/subdimensions	Working groups	X	SD	X	SD	Test value	p
Drug preparation sub-dimension	Jigsaw	21.90	6.06	36.18	2.73	91.974	<0.001*
	Metaverse	21.24	6.18	35.06	3.47		
	Classic learning	21.73	6.00	34.30	4.45		
	Test value	0.991		3.234			
	p	0.374		0.074			
Drug application sub-dimension	Jigsaw	27.04	5.15	36.32	2.72	58.475	<0.001*
	Metaverse	26.12	4.76	33.68	3.25		
	Classic learning	25.08	4.33	34.36	3.72		
	Test value	2.678		4.068			
	p	0.072		0.038*			
	Tukey HSD			1-2			
MASES-C	Jigsaw	49.94	10.09	71.51	4.99	92.995	<0.001*
	Metaverse	52.30	10.01	67.74	6.29		
	Classic learning	46.68	9.70	68.66	7.70		
	Test value	1.956		4.391			
	p	0.145		0.029*			
	Tukey HSD			1-2, 1-3			
*According to the analysis of variance with repeated measurements, it is significant at the level of 0.05, Test value: F test values. SD: Standard deviation, MASES-C: Medication Administration Self-Efficacy Scale in Children							

*According to the analysis of variance with repeated measurements, it is significant at the level of 0.05, Test value: F test values.
SD: Standard deviation, MASES-C: Medication Administration Self-Efficacy Scale in Children

Sensitivity analyses of MASES-C and its preparation and administration dimensions, performed using ANCOVA to estimate differences in least squares means, confirmed the findings of the primary analysis (Table 3).

Students' opinions and suggestions on the educational methods used to gain pediatric drug administration skills: Students' opinions on the evaluation of educational methods were taken with open-ended questions. The statements and suggestions of the students regarding the positive and negative aspects of each education method (jigsaw, metaverse, classic learning) were as follows:

Students' experiences and reflections on jigsaw learning:

Participants in the jigsaw group generally described this method as supportive of learning, enhancing interaction, and well-suited to their own learning styles. Many students emphasized that learning through discussion with peers made the process more enjoyable and memorable. One student reflected, "Learning by discussing with my friends was a fun and memorable method. There was no stress caused by the instructor (A2)." Such statements suggest that jigsaw not only facilitated learning but also created a more relaxed and less stressful environment.

Table 3. Sensitivity analysis using ANCOVA

Scale/subdimensions	Working groups	Pre-test		Test value (F)	p*
		X (95% CI)	SEM		
Drug preparation sub-dimension	Jigsaw	14.28 (12.29-16.27)	0.81	0.605	0.548
	Metaverse	12.82 (10.83-14.81)	1.03		
	Classic learning	13.06 (11.07-15.05)	1.14		
	Difference in the LS mean (95% CI)	-0.46 (-0.53 to -0.40)	0.017		
Drug application sub-dimension	Jigsaw	9.28 (7.66-10.89)	0.72	1.044	0.355
	Metaverse	7.62 (6.01-9.23)	0.86		
	Classic learning	8.30 (6.68-9.91)	0.85		
	Difference in the LS mean (95% CI)	-0.39 (-0.51 to -0.26)	0.016		
MASES-C	Jigsaw	23.56 (20.28-26.84)	1.34	0.933	0.396
	Metaverse	20.44 (17.16-23.72)	1.78		
	Classic learning	21.36 (18.08-24.64)	1.81		
	Difference in the LS mean (95% CI)	-0.46 (-0.61 to -0.30)	0.022		

ANCOVA: Analysis of covariance, CI: Confidence interval, SEM: Standard error of the mean, LS: Least square, *p-value from ANCOVA adjusted for the difference in baseline and post scores

Another recurring theme was the value of interaction. Students noted that working in groups allowed them to recognize their own shortcomings and benefit from one another's perspectives. As one student explained, "It was very good to work as a group, we evaluated our shortcomings together and had the chance to give feedback to each other (A7)." Similarly, several participants highlighted that teaching knowledge to peers reinforced their own understanding and made learning more permanent.

Finally, students felt that the jigsaw method matched their personal learning preferences and should not be limited to this course alone. For example, one student remarked, "I think it's the best way to learn. This method can be used in all courses (A4)." These reflections point to the potential of jigsaw as a flexible approach that can be adapted to different educational contexts.

Taken together, the findings indicate that the jigsaw method provided students with an enjoyable learning experience that reduced anxiety, encouraged collaboration, and supported long-term knowledge retention, while also being perceived as applicable beyond a single course.

Although students in the jigsaw group emphasized many positive aspects, they also pointed out some limitations of the method. A recurring concern was that learning outcomes varied depending on each student's ability to explain the material. As one participant expressed, "I think that learning by yourself and telling your friends is more permanent. However, the topics that my friends told me were less permanent for me (A1)." Similarly, another student noted, "In the jigsaw group, not every individual was at a sufficient level when explaining to their friends. They tried and learned to explain to the other person, but there were those who could not express themselves fully to the other party (A6)." These reflections suggest that while peer teaching can enhance understanding, its effectiveness depends heavily on the communication skills of the group members, which may create disparities in learning.

Another issue raised was that some students did not find the method efficient. For instance, one student simply stated, "I don't think it's an effective method, I didn't get efficiency (A5)." This highlights that although the method was generally well-received, it was not universally effective for all learners.

Alongside these criticisms, students also shared suggestions for improving the course. Interestingly, despite the noted limitations, some participants recommended that the jigsaw method should be used more broadly. As one student explained, "I think it's the best way to learn. This method can be used in all courses (A3)." This indicates that while the method has weaknesses, especially regarding peer expression skills, students still recognized its overall potential and adaptability to different learning contexts.

Students' experiences and reflections on metaverse learning: Students in the metaverse group expressed both positive and negative views about their experiences, along with suggestions for improving the method.

On the positive side, several students emphasized that the metaverse contributed to their learning and was particularly

effective when well organized. One participant explained, "I think that when the metaverse is well managed and well planned, it can be an effective method for our learning. Most of us are now spending time in the virtual world and we are familiar with the virtual environment (B3)." Another added, "I think I learned very well in this class (B1)." These reflections suggest that familiarity with virtual environments may have supported students' adaptation to the method.

Students also appreciated the time-saving aspect of metaverse learning. For instance, one student noted, "I like this method of training better. When I don't understand it, I can watch it over and over again and reinforce the subject (B5)." Another highlighted, "It is very useful for using time effectively (B2)." Such statements indicate that the opportunity to revisit digital content increased flexibility and allowed for self-paced learning.

Despite these positive aspects, students also identified several challenges. A number of participants found the method unproductive or less efficient compared to face-to-face learning. One student explained, "Since I am not very familiar with the virtual world, I found the face-to-face lessons with my lab instructor to be more beneficial to me. Of course, this may vary from person to person (B4)." Technical difficulties were also mentioned, as one student stated, "I had technical problems watching and following some of the videos. I had to watch the video over and over again (B10)." Another reflected, "I don't think it's very efficient, I've learned, but I believe the level remains low. I think it is very difficult to provide learning while taking courses in the metaverse. I have received other training with this method before. Due to the lack of physical materials and environment, it can be difficult to transfer the education received to practice (B9)." Collectively, these views highlight that limitations in technical infrastructure and lack of physical practice opportunities may hinder the effectiveness of metaverse-based education.

Problems with attention and concentration were also reported. For some students, prolonged screen time created difficulties in maintaining focus. As one participant explained, "It wasn't productive for me. I was distracted by looking at the screen trying to learn something (B6)." Another added, "It was bad to have to look at the screen all the time, and the fact that there was no written material to underline affected me badly (B9)." These comments suggest that the absence of tangible learning materials and the reliance on screens may reduce engagement for certain learners.

Alongside these criticisms, students proposed suggestions to improve the method. Some highlighted the need to combine the metaverse with other approaches, as one student expressed, "I don't think it was very efficient because it was a method I wasn't familiar with, but it was a different experience. If it is combined with other learning methods and developed a little more, good results can be obtained (B8)." Others emphasized its potential for broader application, noting, "A good method for learning, can be applied to all subjects." Finally, one participant suggested increasing the available materials, stating, "The learning technique is good. There may be more videos (B7)."

Taken together, these findings indicate that while students recognized the potential of the metaverse to support learning and time efficiency, they also highlighted challenges such as technical issues, attention difficulties, and the absence of physical practice materials. Importantly, their suggestions for integration with other methods and expansion of resources point to ways in which the effectiveness of metaverse learning could be enhanced in the future.

Students' experiences and reflections on classic learning: Participants in the classic learning group expressed both positive and negative views regarding their experiences. On the positive side, students described this method as supportive of learning, aligned with their personal preferences, and providing opportunities for interaction. Several participants highlighted its familiarity and usefulness: "It was very good, I think it was useful" (C2); "Previously, I had the chance to take courses in different courses with different teaching methods. However, in my opinion, I think it is the most effective learning method" (C10). Another student emphasized its routine nature, stating, "I did not experience a negative situation because this learning model is a method that we usually use and that I am familiar with" (C1). Similarly, one participant noted, "The classic learning method is always my preference" (C7).

Beyond familiarity, the interactive nature of face-to-face instruction was valued. As one student reflected, "It is a more effective method for me to teach an interactive lesson with the instructor, to immediately ask about a part I do not understand and to reinforce it instantly. In addition, receiving face-to-face training, touching, and applying training materials provides more permanent learning" (C4). Such statements suggest that classic learning was appreciated for its structured delivery and the opportunity to engage directly with instructors and materials.

Despite these positive reflections, participants also identified limitations of the classic method. A recurring concern was attention and concentration problems. For example, one student explained, "I think that the classic learning method is a type of lesson in which the student is less active, which usually proceeds in one direction and bores the student after a certain period of time. The student is unable to concentrate their attention and interest in this method" (C3).

Another limitation mentioned was the heavy reliance on the instructor's expression skills. As one participant stated, "We didn't have a lot of participation in the class, I think monologue is a learning method. Our learning skills change according to the instructor who tells us. It is a method that is explained only by the instructor according to himself, and our opinions are not taken" (C9). These reflections indicate that the effectiveness of classic learning may vary depending on the instructor's style and level of engagement.

Students also noted that the method could be time-consuming. One student remarked, "I can learn, but it takes a lot of time. This method can be used in lessons that are in the form of direct lectures. I think it can be useful if it is enriched with other educational methods. Otherwise, listening to someone for a long time is both boring and negatively affects

learning" (C6). Another participant added, "It could have been improved in better and more engaging ways. The classic learning method reduces interest and focus on the lesson. I think the energy and attention of the instructor decreases as the time progresses" (C8).

Taken together, these reflections show that while students valued the familiarity, structure, and interactive elements of classic learning, they also pointed out its limitations related to attention, reliance on the instructor's expression skills, and the risk of monotony over time. Suggestions for improvement emphasized enriching the approach with diverse educational strategies to make the learning process more engaging and effective.

DISCUSSION

Preventing medication errors is crucial for pediatric patient safety. Continuing education for nurses and integrating innovative learning methods in undergraduate training enhance drug administration quality and safety.^{33,34} This study assessed nursing students' self-efficacy in pediatric drug administration using three learning methods, revealing significant differences among them.

In our study, nursing students in the jigsaw group had higher self-efficacy in pediatric drug administration than those in the metaverse and classic learning groups. Similarly, Balkaya et al.⁷ found that jigsaw learning improved nursing students' motivation and self-confidence. Yang et al.³⁵ also reported that cooperative learning enhanced students' motivation, attitude, problem-solving confidence, and pediatric nursing knowledge. These findings suggest that cooperative and interactive strategies provide not only cognitive but also affective benefits, which are crucial for nursing students who often face stress and performance pressure in clinical environments. In our study, students were asked about their opinions on the three learning methods. The majority preferred the jigsaw technique and believed that combining it with classic learning could enhance both theoretical and clinical practice skills. This indicates that students perceive blended approaches as more effective than relying on a single method, which aligns with current trends in nursing education emphasizing hybrid and student-centered pedagogies. This perspective may stem from the increasing number of nursing students, the shortage of faculty compared to student numbers, and the effective outcomes achieved by integrating innovative methods like the jigsaw technique with traditional learning.

In today's nursing education, it is crucial for students to transform theoretical knowledge into practical skills, as nursing integrates theory, communication, clinical practice, and empathy.^{36,37} To prevent pediatric medication errors and fulfill hospital responsibilities, nurses must develop continuous learning and intellectual skills. Thus, active learning methods are preferred over memorization-based education.^{38,39} Jeppu et al.⁴⁰ found that the jigsaw technique significantly enhances academic achievement and psychomotor skills compared to traditional learning. Our study also revealed that students favored this method due to its stress-free, interactive nature, allowing open communication

and collaboration. Clinical stress often arises from instructor criticism and fear of mistakes.⁴¹ The jigsaw technique fosters peer support, reducing anxiety and improving engagement, making it an effective learning approach for nursing students. This is particularly relevant in pediatric drug administration, where both accuracy and confidence are essential for patient safety.

The use of the metaverse has increased during extraordinary events like pandemics and earthquakes, providing virtual interaction opportunities. Studies suggest it positively impacts education by maintaining engagement.¹⁰ In our study, students reported technical issues, distractions from avatars, and difficulty focusing on screens for long periods. They frequently rewound presentations for clarity, suggesting metaverse learning is more effective when combined with other methods. However, they appreciated the ability to rewatch recorded sessions. Similar studies present mixed results on its effectiveness. While Lee et al.¹⁴ found meta-scenarios beneficial, De Gagne et al.¹² noted both advantages and challenges, and Plotzky et al.¹⁵ emphasized the need for social interaction. These inconsistencies indicate that metaverse learning is still in its early adoption phase, and its effectiveness largely depends on design quality, technological infrastructure, and integration with other pedagogical strategies. Our findings confirm that although metaverse platforms provide flexibility and continuity during crises, they are not yet sufficient to replace face-to-face or interactive peer learning methods. Regarding the drug preparation sub-dimension, although the jigsaw group showed higher post-test scores, this difference did not reach statistical significance ($p=0.074$). Our research found that classic learning was more effective than metaverse-based learning in enhancing self-efficacy in pediatric drug administration. Classic learning was familiar, useful, and allowed interactive lessons. However, its drawbacks included one-directional teaching, boredom, decreased attention, varying comprehension levels, time consumption, instructor fatigue, stress from instructors, and limited student expression. This duality reflects a broader debate in nursing education: while traditional methods ensure structure and familiarity, they may hinder active engagement and deeper learning unless enriched with complementary approaches.

Students suggested integrating active learning methods to enhance classic learning. Similarly, Mutlu et al.⁴² found that nursing students' anxiety negatively affected their self-efficacy in child health courses. Studies show that nursing students experience the highest stress during pediatrics internships.^{43,44} Unlike classic learning, jigsaw learning fosters group responsibility, shared leadership, and educator supervision, promoting commitment and self-confidence.⁴⁵ Innovative methods may help reduce anxiety and enhance clinical self-efficacy. Medical errors, particularly in drug administration, are a major global challenge, increasing costs, morbidity, and mortality.³ To prevent such errors, integrating theoretical knowledge with clinical practice is essential in nursing education.² Therefore, our results emphasize that relying solely on classical approaches may not be sufficient to prepare students for the complex, high-stakes tasks of

pediatric nursing, and hybrid models that include cooperative learning and emerging technologies are needed.

Recommendations

Future research should be conducted across multiple institutions and larger sample groups to strengthen external validity. Longitudinal studies could explore the sustainability of self-efficacy gains over time. Additionally, experimental studies comparing blended learning models combining jigsaw, metaverse, and classic approaches-would clarify the most effective strategies for pediatric drug administration training.

Implications for practice

For nurse educators, these findings highlight the importance of diversifying teaching strategies. While jigsaw learning fosters collaboration and confidence, metaverse platforms can provide continuity during disruptions, and classic learning ensures structure. A blended model incorporating all three methods may optimize student engagement and preparedness. In practice, adopting such an integrated approach could contribute to reducing pediatric medication errors, improving patient safety, and enhancing nursing students' readiness for clinical responsibilities.

Limitations

A strength of this study is its mixed-methods design, which provided both quantitative and qualitative perspectives on student self-efficacy. This approach allowed us to capture not only measurable outcomes but also students' reflections, adding depth to the findings. However, limitations should also be acknowledged. The study was conducted in a single institution with a limited sample size, which may restrict generalizability. In addition, the study measured only self-efficacy and not actual clinical performance; therefore, the results reflect students' perceived competence rather than directly observed skills. Furthermore, self-reported measures may not always reflect actual clinical performance, and qualitative findings may be influenced by participants' willingness to share their experiences.

CONCLUSION

Our research shows that jigsaw, metaverse, and classic learning methods can support pediatric nursing students' self-efficacy in drug administration to varying degrees. The jigsaw method was the most effective, yielding the highest self-efficacy scores due to its interactive environment and stress reduction. The metaverse provided advantages such as time management and flexibility, but its overall effectiveness was limited by technical issues and distractions, resulting in lower scores compared to the other methods. The classic method allowed for direct instructor interaction but was perceived as monotonous and time-consuming. Developing hybrid models by combining the strengths of each method could improve educational efficiency. Future research should explore these hybrid approaches to develop more holistic and flexible educational strategies.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Süleyman Demirel University Faculty of Medicine Clinical Researches Ethics Committee (Date: 13.04.2023, Decision No: 87)

Informed Consent

Written informed consent was obtained from all students who participated in the study.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Howlett MM, Cleary BJ, Breatnach CV. Defining electronic-prescribing and infusion-related medication errors in paediatric intensive care: a Delphi study. *BMC Med Inform Decis Mak*. 2018;18(1):130. doi:10.1186/s12911-018-0713-8
- Alghamdi AA, Keers RN, Sutherland A, Ashcroft DM. Prevalence and nature of medication errors and preventable adverse drug events in paediatric and neonatal intensive care settings: a systematic review. *Drug Saf*. 2019;42(12):1423-1436. doi:10.1007/s40264-019-00856-9
- Izadpanah F, Nikfar S, Bakhshi Imcheh F, Amini M, Zargaran M. Assessment of frequency and causes of medication errors in pediatrics and emergency wards of teaching hospitals affiliated to Tehran University of Medical Sciences (24 hospitals). *J Med Life*. 2018;11(4):299-305. doi:10.25122/jml-2018-0046
- D'Errico S, Zanon M, Radaelli D, et al. Medication errors in pediatrics: Proposals to improve the quality and safety of care through clinical risk management. *Front Med (Lausanne)*. 2022;8:814100. doi:10.3389/fmed.2021.814100
- Ziyai NY, Bozkurt R, Kilickiran H, Dogu O. The effect of flipped-jigsaw learning models on ethical decision-making. *Nurs Ethics*. 2024;31(2-3):132-147. doi:10.1177/09697330231187139
- Çetinkaya Uslusoy E, Aydınli A, Durna F. Enhancing learning motivation and academic achievement in nursing students through metaverse-based learning: a randomized controlled study. *Jpn J Nurs Sci*. 2024;21(3):e12594. doi:10.1111/jjns.12594
- Balkaya NA, Bilge C, Dönmez S. The effect of cooperative learning and interactive education methods on the motivation and self-confidence of nursing students. *J Women's Health Environ Nutr*. 2022;8(1):14-24.
- Ongor M, Uslusoy EC. The effect of multimedia-based education in e-learning on nursing students' academic success and motivation: a randomised controlled study. *Nurse Educ Pract*. 2023;71:103686. doi:10.1016/j.nepr.2023.103686
- Goolsarran N, Hamo CE, Lu WH. Using the jigsaw technique to teach patient safety. *Med Educ Online*. 2020;25(1):1710325. doi:10.1080/10872981.2019.1710325
- Jeon JH. A study on education utilizing metaverse for effective communication in a convergence subject. *Int J Internet Broadcast Commun*. 2021;13(4):129-134.
- Moon H, Freidouny M, Rajabi MS, Bozorgmehrian S, Sangwan A, Jeon M. The influence of olfactory and visual stimuli on students' performance and mood in virtual reality environment. *Proc Hum Factors Ergon Soc Annu Meet*. 2023;67(1):2441-2446.
- De Gagne JC, Randall PS, Rushton S, et al. The use of metaverse in nursing education: an umbrella review. *Nurse Educ*. 2023;48(3):E73-E78. doi:10.1097/NNE.0000000000001327
- Bryant R, Vogt M, Miller C. Effects of an avatar-based simulation on family nurse practitioner students' self-evaluated suicide prevention knowledge and confidence. *Nurs Educ Perspect*. 2024. doi:10.1097/01.NEP.0000000000001321
- Lee H, Yi Y, Moon W, Yeo JY. Exploring the potential use of the metaverse in nurse education through a mock trial. *Nurse Educ Today*. 2023;131:105974. doi:10.1016/j.nedt.2023.105974
- Plotzky C, Lindwedel U, Sorber M, et al. Virtual reality simulations in nurse education: a systematic mapping review. *Nurse Educ Today*. 2021;101:104868. doi:10.1016/j.nedt.2021.104868
- Cho IY, Yun JY, Moon SH. Development and effectiveness of a metaverse reality-based family-centered handoff education program in nursing students. *J Pediatr Nurs*. 2024;76:176-191. doi:10.1016/j.pedn.2024.02.005
- Kim DY, Lee HK, Chung K. Avatar-mediated experience in the metaverse: the impact of avatar realism on user-avatar relationship. *J Retail Consum Serv*. 2023;73:103382.
- Cameron R. Lessons from the field: Applying the good reporting of a mixed methods study (GRAMMS) framework. *Electron J Bus Res Methods*. 2013;11(2):55-66.
- Cuschieri S. The CONSORT statement. *Saudi J Anaesth*. 2019;13(Suppl 1):S27-S30.
- Du KJ, Li GS, Zhang K, Lin Y, Yang F, Hannes K. Prof. Karin Hannes: COREQ (Consolidated Criteria for Reporting Qualitative Studies). *Ann Transl Med*. 2022;10(19):1073. doi:10.21037/atm-2022-23
- Guest G, Bunce A, Johnson L. How many interviews are enough? An experiment with data saturation and variability. *Field Methods*. 2006;18(1):59-82. doi:10.1177/1525822X05279903
- Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol*. 2006;3(2):77-101. doi:10.1191/1478088706qp0630a
- Bektaş İ, Yardımcı F, Bektaş M. Development of medication administration self-efficacy scale in children for nursing students and psychometric properties. *DEUHFED*. 2021;14(4):320-327.
- Spatial.io. Creating collaborative learning environments in the metaverse. Spatial.io Reports. Accessed March 1, 2023. <https://www.spatial.io/>
- Sezer H, Gül A. Health sciences students' perspectives on metaverse and digital technology use: cross-sectional descriptive study. *BMC Nurs*. 2024;23:656. doi:10.1186/s12912-024-02309-w
- Sendra-Portero F, Lorenzo-Álvarez R, Rudolphi-Solero T, Ruiz-Gómez MJ. The Second Life metaverse and its usefulness in medical education after a quarter of a century. *J Med Internet Res*. 2024;26:e59005. doi:10.2196/59005
- Ghaempanah F, Alizadeh M, Hosseini R, et al. Metaverse and its impact on medical education and health care system: a narrative review. *Health Sci Rep*. 2024;7:e70100. doi:10.1002/hsr2.70100
- Unity Technologies. VR and AR applications in education. Unity Whitepaper. Accessed March 1, 2023. <https://unity.com/>
- Ben Yahya L, Naciri A, Radid M, Chemsı G. Immersive simulation in nursing and midwifery education: a systematic review. *J Educ Eval Health Prof*. 2024;21:19. doi:10.3352/jeehp.2024.21.19
- Park S, Shin HJ, Kwak H, Lee HJ. Effects of immersive technology-based education for undergraduate nursing students: Systematic review and meta-analysis (GRADE). *J Med Internet Res*. 2024;26:e57566. doi:10.2196/57566
- Amazon Web Services (AWS). Cloud computing for education: Scaling digital classrooms. AWS Educational Resources. Accessed March 1, 2023. <https://aws.amazon.com/education/>
- Wondmieneh A, Alemu W, Tadele N, Demis A. Medication administration errors and contributing factors among nurses: a cross-sectional study in tertiary hospitals, Addis Ababa, Ethiopia. *BMC Nurs*. 2020;19:4. doi:10.1186/s12912-020-0397-0
- Sachdeva S, Bhatia S, Al Harrasi A, et al. Unraveling the role of cloud computing in health care system and biomedical sciences. *Heliyon*. 2024;10(7):e29044. doi:10.1016/j.heliyon.2024.e29044
- Hockenberry MJ, Bryant R, Hellsten MB. Wong's Clinical Manual of Pediatric Nursing. 9th ed. St. Louis, MO: Elsevier Health Sciences; 2023.
- Yang BH, Chung CY, Li YS, Lu CF. A cooperative learning intervention for improving a simulation-based paediatric nursing course: a quasi-experimental study. *Nurse Educ Pract*. 2024;80:104149. doi:10.1016/j.nepr.2024.104149

36. Fields L, Trostian B, Moroney T, Dean BA. Active learning pedagogy transformation: a whole-of-school approach to person-centred teaching and nursing graduates. *Nurse Educ Pract.* 2021;53:103051. doi:10.1016/j.nepr.2021.103051
37. Gassas R. Sources of the knowledge-practice gap in nursing: lessons from an integrative review. *Nurse Educ Today.* 2021;106:105095. doi:10.1016/j.nedt.2021.105095
38. Lin C, Han C, Huang Y, Chen L, Su CC. Effectiveness of the use of concept maps and simulated cases as a teaching-learning strategy in enhancing the learning confidence of baccalaureate nursing students: a qualitative approach. *Nurse Educ Today.* 2022;115:105418. doi:10.1016/j.nedt.2022.105418
39. Ryu H, Lee H, Yoo HJ. Development of a metaverse online learning system for undergraduate nursing students: a pilot study. *Nurse Educ.* 2024;49(2):E74-E79. doi:10.1097/NNE.0000000000001509
40. Jeppu AK, Kumar KA, Sethi A. We work together as a group: implications of jigsaw cooperative learning. *BMC Med Educ.* 2023;23(1):734. doi:10.1186/s12909-023-04734-y
41. Al-Sayaghi KM, Atrous MH, Alkaluby EA, et al. Relationship between self-esteem and stress coping strategies among internship nursing students. *Open Access Maced J Med Sci.* 2023;11(G):26-32.
42. Mutlu B, Azak M, Kalaycı S. Investigation of the relationship between clinical comfort-worry and self-efficacy perceptions of nursing students who have taken pediatric nursing course. *OPUS Int J Soc Res.* 2020;15(26):4293-4315.
43. Lassche M, Al-Qaaydeh S, Macintosh CI, Black M. Identifying changes in comfort and worry among pediatric nursing students following clinical rotations. *J Pediatr Nurs.* 2013;28(1):48-54. doi:10.1016/j.pedn.2011.12.008
44. Kubin L, Wilson CE. Effects of community volunteer children on student pediatric assessment behaviors. *Clin Simul Nurs.* 2017;13(7):303-308.
45. Abdel-Mordy MA, Sabry SS, Abdelrazek A. Effect of cooperative jigsaw learning strategy on community nursing students' attitude and achievement. *Int Egypt J Nurs Sci Res.* 2022;2(2):487-500.

The efficacy of transperineal and transabdominal ultrasound findings in predicting spontaneous vaginal delivery during active labor

✉Ekin Büyüknisan, ✉Neşe Yücel, ✉Oğuz Yücel, ✉Şule Gül Aydın, ✉Esra Yazkan

Department of Obstetrics and Gynecology, Adana City Hospital, Adana, Türkiye

Cite this article as: Büyüknisan E, Yücel N, Yücel O, Aydın ŞG, Yazkan E. The efficacy of transperineal and transabdominal ultrasound findings in predicting spontaneous vaginal delivery during active labor. *J Health Sci Med.* 2025;8(6):1076-1081.

Received: 14.09.2025

Accepted: 09.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This research sought to assess the clinical value of transperineal and transabdominal ultrasound measurements, particularly angle of progression (AoP), head-perineum distance (HPD), and midline angle (MLA), in relation to mode of delivery during the active phase of labor.

Methods: In this prospective observational design, pregnant women in active labor were enrolled and underwent repeated transperineal and transabdominal ultrasound examinations. Hourly AoP, HPD and MLA measurements were performed. Patients were classified according to delivery mode (spontaneous vaginal birth or cesarean section) and the changes in ultrasound parameters were compared between groups. Multiparous and nulliparous women were also analyzed separately.

Results: Mean AoP values remained significantly greater in the vaginal delivery group throughout all time points when compared to the cesarean group ($p<0.001$). HPD measurements were significantly increased in the cesarean group beginning as early as the first hour ($p\leq0.001$). Multiparous women had consistently higher AoP values than nulliparous women throughout active labor ($p<0.01$), with the difference persisting at the onset of the second stage. Analysis revealed that smaller AoP values at the onset of active labor were significantly related to longer labor courses, confirming a negative correlation. Higher MLA measurements were observed in the cesarean group within the first hours of labor ($p<0.001$).

Conclusion: Serial measurements of AoP, HPD and MLA by ultrasound provide objective and reliable information for predicting delivery mode in active labor. By integrating these ultrasound findings into practice, clinicians may be able to make better decisions and manage labor more efficiently.

Keywords: Fetal head, labor, obstetric, ultrasonography, prenatal

INTRODUCTION

Accurate and timely intervention in the management of labor is crucial for both maternal and neonatal health.¹ Objective assessment of labor progression remains a frequently debated topic in clinical practice. Although vaginal examination is traditionally regarded as the gold standard for intrapartum monitoring, this method is limited by its subjectivity and the increased risk of infection.^{2,3}

Ultrasonography has become an increasingly common method for objectively assessing labor in contemporary obstetric practice. Transperineal and transabdominal ultrasound techniques offer significant advantages, particularly in evaluating parameters such as fetal head position, progression, and its location within the pelvic cavity.^{4,5} Data obtained from ultrasonography provide clinicians with valuable information for predicting labor progression and the mode of delivery.

The literature contains a wide range of studies examining the role of transperineal and transabdominal ultrasound in monitoring labor and forecasting spontaneous vaginal delivery.^{6,7} However, there is still no clear consensus regarding

the effectiveness of these methods and their integration into routine clinical practice for women in active labor. In particular, the comparison of different ultrasonographic parameters and the identification of the most reliable indicators remain central topics of current research.⁸

The purpose of this study was to assess how transperineal and transabdominal ultrasound findings can be applied to monitor labor progression and predict spontaneous vaginal delivery in women during active labor. In this way, we sought to contribute to the development of more objective and reliable methods for labor management.

METHODS

This research was designed as a prospective observational study. Ethics approval was obtained from the Adana City Hospital Scientific Researches Ethics Committee (Date: 25.07.2024, Decision No: 76). All participants provided informed consent before being enrolled in the study. All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Corresponding Author: Şule Gül Aydın, sulegulaydin@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

Sample size and power analysis was performed based on previous literature indicating a 10-12° difference in the angle of progression (AoP) between delivery modes (SD $\approx$ 9-10°). A minimum of 60 participants was estimated to achieve 80% power at $\alpha=0.05$. A total of 150 women were enrolled. Post-hoc power analysis using the observed baseline AoP difference (100.6 $\pm$ 7.4 vs 88.1 $\pm$ 9.1) demonstrated a power greater than 99%.

The study population consisted of 150 pregnant women in active labor, between 34+0 and 40+6 weeks of gestation, who presented to the Department of Obstetrics and Gynecology at Adana City Hospital between August 1, 2024 and May 15, 2025. Of these participants, 65 were nulliparous and 85 were multiparous. The study population was restricted to cases meeting the following criteria: singleton gestation, cephalic presentation, maternal age of 18-45 years and active labor status. Exclusion criteria included multiple pregnancy, a history of major uterine or cervical surgery, fetal anomaly or malpresentation, previous cesarean delivery, body-mass index (BMI) over 35, and age under 18 or over 45 years. Women requiring intrapartum magnesium sulfate or nifedipine, or those with preeclampsia, eclampsia, or other severe obstetric complications at admission, were not included in the study. No participant received intrapartum magnesium sulfate or nifedipine during data collection. Cases with gestational hypertension controlled with oral methyldopa were included.

All participants included in the study were monitored at regular intervals throughout labor after hospital admission. Age, height, weight, and body-mass index were among the baseline demographic variables collected for all participants. Obstetric history, including gravida, parity, abortion, and ectopic pregnancy, was documented. Existing chronic diseases (e.g., gestational hypertension, gestational diabetes, hypothyroidism), prior surgeries, and current medications were also recorded in detail. Gestational age was defined by maternal last menstrual period and validated through early pregnancy ultrasound, in accordance with the study's inclusion requirements. Information on delivery mode (spontaneous vaginal or cesarean), cesarean indications, labor duration, and the onset of the second stage was recorded systematically. All these data were incorporated into the dataset for analysis along with ultrasonographic and digital pelvic examination findings.

Ultrasonographic examinations were performed using a Samsung V8 color Doppler ultrasonography system equipped with a convex probe (Samsung Medison Co., Ltd., South Korea), and conducted by the clinician responsible for the study. All ultrasound assessments were carried out by Dr. Ekin Büyüknisan, a fourth-year obstetrics and gynecology resident with substantial experience in intrapartum ultrasonography. Conducting all scans by a single operator ensured measurement consistency and minimized inter-observer variability. Prior to assessment, all patients were placed in the dorsal lithotomy position, and their bladders were emptied using a sterile 16 F Foley catheter. To optimize image quality during transperineal ultrasonography, ample ultrasound gel was applied between the skin and the probe. Obstetric assessment included evaluation of gestational age, fetal growth, fetal

heart rate, presentation, and placental position. Assessment of gestational age relied on the maternal last menstrual period, with confirmation provided by early pregnancy ultrasound. According to ISUOG recommendations, the AoP corresponds to the angle created by a line extending along the maternal symphysis pubis (from superior to inferior) and another line drawn tangentially from the inferior border of the symphysis toward the fetal occiput (Figure 1). After obtaining sagittal plane images, the ultrasound probe was rotated into the transverse plane. With gentle pressure applied toward the pubic bone to avoid patient discomfort, the midline angle (MLA) was calculated as the angle between the echogenic line of the interhemispheric fissure and the anteroposterior axis of the symphysis pubis (Figure 2). We defined head-perineum distance as the least distance observed between the most protruding contour of the fetal head and the surface of the perineum (Figure 3).⁹

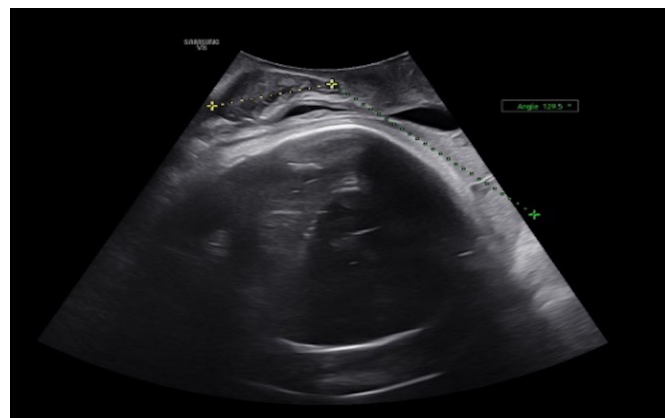


Figure 1. Intrapartum transperineal ultrasound. Measurement of angle of progression



Figure 2. Intrapartum transperineal ultrasound. Measurement of midline angle

All patients in active labor underwent both transabdominal and transperineal ultrasonography, as well as digital pelvic examination, at hourly intervals once cervical dilation reached 6 cm or more. Findings from each method were systematically recorded. Digital pelvic examination assessed cervical dilation, fetal head station, and cervical effacement, and served as the traditional reference for monitoring labor progression. Transabdominal USG was used to objectively evaluate the fetal head position as occiput anterior, posterior, or transverse. With transperineal USG, the following parameters



Figure 3. Intrapartum transperineal ultrasound. Measurement of head-perineum distance

were measured: fetal head station, angle of progression (AoP), head-perineum distance (HPD), MLA, head-symphysis distance (HSD), and cervical dilation. These objective measurements were analyzed to assess labor progression and predict the mode of delivery. Additionally, total labor duration was calculated in minutes for each patient and recorded along with the mode of delivery. All data obtained were used for intergroup comparisons and predictive analyses.

Statistical Analysis

Data were analyzed using IBM SPSS Statistics for Windows, Version 24.0 (IBM Corp., Armonk, NY, USA). Descriptive statistics were expressed as mean±standard deviation (SD) or number (percentage), as appropriate. The normality of distribution was assessed using the Kolmogorov–Smirnov test. Comparisons between two independent groups were performed using the independent samples t-test or the Mann-Whitney U test, depending on data distribution. Chi-square (χ^2) test or Fisher's exact test was used for categorical variables. Correlations between continuous variables were evaluated using Pearson's correlation coefficient. A p-value <0.05 was considered statistically significant.

RESULTS

The baseline demographic and clinical characteristics of the study population are summarized in **Table 1**. The mean maternal age was 26.1±4.2 years, with an average BMI of 29.3±4.7 kg/m² and a mean gestational age of 38.9±1.5 weeks at delivery. Among all participants, five women (3.3%) delivered preterm between 34+0 and 36+6 weeks (three spontaneous vaginal births and two cesarean sections). Because the distribution of preterm cases was comparable between groups, no statistically significant difference in the main outcomes was observed. Of the 150 women, 43.3% were nulliparous and 56.7% multiparous. Most participants had no history of chronic disease (85.3%) or surgery (96.7%), and the majority were not on medication during pregnancy (89.3%). Spontaneous vaginal delivery occurred in 89.3% of cases, while 10.7% required cesarean section, primarily due to arrest of labor (81.3%).

At the onset of the active phase (hour 0), women who ultimately delivered vaginally exhibited significantly greater average AoP values than those who underwent cesarean section. (100.6±7.4 vs. 88.1±9.1, $p<0.001$). The gap in AoP values between the two

Table 1. Demographic and clinical characteristics of the study population

Characteristic	Mean±SD/n (%)
Age (years)	26.1±4.2
Weight (kg)	79.0±12.5
Height (cm)	163.7±10.6
BMI (kg/m ²)	29.3±4.7
Gravida	2.2±1.1
Parity (mean)	1.0±1.0
Abortus	0.3±0.6
Gestational age (weeks)	38.9±1.5
Duration of labor (min)	150.8±53.3
Parity group	Nulliparous: 65 (43.3%)/multiparous: 85 (56.7%)
History of ectopic pregnancy	No: 145 (96.7%)/yes: 5 (3.3%)
Chronic disease	No: 128 (85.3%)/Gest. HT: 9 (6.0%)/Gest. GDM: 6 (4.0%)/hypothyroidism: 7 (4.7%)
History of surgery	No: 145 (96.7%)/yes: 5 (3.3%)
Medications used	None: 134 (89.3%)/methyldopa: 6 (4.0%) / Multiple drugs: 10 (6.7%)
Mode of delivery	SVD: 134 (89.3%)/C/S: 16 (10.7%)
Indication for C/S	Arrest of labor: 13 (81.3%)/AFD: 3 (18.8%)

SD: Standard deviation, BMI: Body-mass index, HT: Gestational hypertension, GDM: Gestational diabetes mellitus, SVD: Spontaneous vaginal delivery, C/S: Cesarean section

groups remained significant at every subsequent time point, with the vaginal delivery group consistently showing higher mean values (all $p<0.001$). The difference was especially pronounced at the 5th hour, where the AoP reached 145.2±4.0° in the vaginal birth group, while remaining markedly lower in the C/S group (97.0°) (**Table 2**).

Table 2. Hourly mean AoP values according to the mode of delivery in pregnant women in active labor

	SVD (n=134) Mean±SD	C/S (n=16) Mean±SD	p-value
AoP baseline (onset of active labor)	100.6±7.4	88.1±9.1	<0.001
AoP 1 st hour (°)	112.9±11.2	91.2±7.9	<0.001
AoP 2 nd hour (°)	122.0±10.5	93.8±6.7	<0.001
AoP 3 rd hour (°)	133.9±11.3	95.5±7.1	<0.001
AoP 4 th hour (°)	143.6±6.6	95.0±8.2	<0.001
AoP 5 th hour (°)	145.2±4.0	97.0±0.0	<0.001

Statistically significant differences are indicated by bold p-values. Baseline refers to the onset of the active phase of labor. AoP: Angle of progression, SVD: Spontaneous vaginal delivery, C/S: Cesarean section, SD: Standard deviation

At the beginning of the active phase (hour 0), HPD values did not differ significantly between the vaginal delivery and cesarean section groups. (46.3±8.0 mm vs. 46.0±6.8 mm, $p=0.896$). Beginning with the first hour, HPD values were noticeably higher in the cesarean section group than in the vaginal birth group with differences reaching statistical significance at each subsequent hourly measurement ($p\leq 0.001$ for all). The most striking disparity was observed at the 5th hour, where the mean HPD in the C/S group remained markedly elevated (45.0±2.4 mm), while it progressively decreased in the vaginal birth group (16.9±4.1 mm, $p<0.001$) (**Table 3**).

Table 3. Comparison of hourly HPD values according to mode of delivery

	SVD (n=134) Mean±SD	C/S (n=16) Mean±SD	p-value
HPD Baseline (mm)	46.3±8.0	46.0±6.8	0.896
HPD 1 st hour (mm)	34.2±11.9	44.1±6.9	0.001
HPD 2 nd hour (mm)	29.2±8.7	42.0±6.8	<0.001
HPD 3 rd hour (mm)	24.9±7.3	39.9±7.6	<0.001
HPD 4 th hour (mm)	18.6 ± 6.1	37.7 ± 3.5	<0.001
HPD 5 th hour (mm)	16.9 ± 4.1	45.0 ± 2.4	<0.001

Values are presented as mean±SD. Bold p-values indicate statistical significance. Baseline refers to the onset of the active phase of labor. HPD: Head-perineum distance, SVD: Spontaneous vaginal delivery, C/S: Cesarean section, SD: Standard deviation

At hour 0, corresponding to the onset of active labor, multiparous women exhibited significantly elevated mean AoP compared with their nulliparous counterparts (102.7±7.2 vs. 97.8±6.7 $p<0.001$). This significant difference continued through the first four hours of active labor, with multiparous women consistently exhibiting greater AoP values than their nulliparous counterparts (all $p<0.01$). Although the AoP values converged at the fifth hour (145.3±3.8 vs. 145.0±7.1, $p=0.909$) at the beginning of the second stage of labor, multiparous women still showed a statistically higher AoP (144.2±7.1 vs. 142.5±6.5, $p=0.037$) (Table 4).

Table 4. Comparison of hourly mean AoP values in nulliparous and multiparous pregnant women undergoing spontaneous vaginal delivery

	Nulliparous (n=57) Mean±SD	Multiparous (n=77) Mean±SD	p-value
AoP baseline (°)	97.8±6.7	102.7±7.2	<0.001
AoP 1 st hour (°)	107.7±7.1	116.7±12.2	<0.001
AoP 2 nd hour (°)	118.5±8.7	125.2±11.1	<0.001
AoP 3 rd hour (°)	130.7±10.5	137.1±11.2	0.003
AoP 4 th hour (°)	142.0±7.2	146.4±3.9	0.008
AoP 5 th hour (°)	145.3±3.8	145.0±7.1	0.909
AoP at the onset of the second stage (°)	142.5±6.5	144.2±7.1	0.037

Bold p-values indicate statistical significance. Baseline is defined as the onset of active labor. Values are presented as mean±SD. AoP: Angle of progression, SD: Standard deviation

A weak but significant correlation was observed between AoP at the onset of the active phase (0th hour) and labor duration ($r=-0.168$, $p=0.042$), indicating that lower initial AoP values are associated with longer labor. This association became even more pronounced at the 1st hour, where a stronger negative correlation was detected ($r=-0.338$, $p<0.001$) (Table 5).

At hour 0, corresponding to the beginning of the active phase, the cesarean section group showed markedly higher MLA values than the vaginal birth group. (60.0±14.9 vs. 28.8±7.9, $p<0.001$). This marked difference persisted through the first four hours of labor, with the C/S group consistently exhibiting substantially higher MLA values than the vaginal birth group at each time point (all $p<0.001$). By the 5th hour, although the MLA remained higher in the C/S group, the difference was no longer statistically significant (20.0±0.0 vs. 5.7±5.5, $p=0.204$) (Table 6).

Table 5. Correlation between AoP values and duration of labor in pregnant women undergoing spontaneous vaginal delivery

	Pearson correlation	Sig. (2-tailed)	n
AoP baseline (°)	-0.168	0.042	134
AoP 1 st hour (°)	-0.338	<0.001	134
AoP 2 nd hour (°)	-0.059	0.518	121
AoP 3 rd hour (°)	0.187	0.058	105
AoP 4 th hour (°)	-0.085	0.505	64
AoP 5 th hour (°)	0.148	0.645	12

Values are presented as Pearson correlation coefficients and p-values [Sig. (2-tailed)]. Bold p-values indicate statistical significance. Baseline is defined as the onset of active labor. AoP: Angle of progression

Table 6. Comparison of hourly MLA values according to mode of delivery

	SVD (n=134) Mean±SD	C/S (n=16) Mean±SD	p-value
MLA baseline (°)	28.8±7.9	60.0±14.9	<0.001
MLA 1 st hour (°)	21.5±10.5	55.1±13.9	<0.001
MLA 2 nd hour (°)	15.2±9.1	48.2±15.0	<0.001
MLA 3 rd hour (°)	9.3±6.2	38.9±15.2	<0.001
MLA 4 th hour (°)	6.7±5.0	31.5±7.2	<0.001
MLA 5 th hour (°)	5.7±5.5	20.0±0.0	0.204

Values are presented as mean±SD. Bold p-values indicate statistical significance. Baseline is defined as the onset of active labor. MLA: Midline angle, SVD: Spontaneous vaginal delivery, C/S: Cesarean section, SD: Standard deviation

DISCUSSION

In our study, it was determined that AoP values measured by transperineal ultrasound during active labor are a useful parameter in assessing the mode of delivery, and that hourly increases in AoP were significantly greater in women delivering vaginally compared to the cesarean group. AoP increase was more pronounced in the vaginal delivery group, especially in the last hours of the active phase, while progression remained limited in the cesarean section group.

The relatively small number of women in the cesarean section group (n=16) may have limited the statistical power of our analyses. Nevertheless, our findings suggest that an early intrapartum AoP value below 100° may be a clinically meaningful indicator for cesarean delivery. Therefore, routinely including AoP assessment during the active phase of labor may facilitate earlier identification of women at high risk for surgical intervention and support more evidence-based intrapartum management.

Bulut and Ceyhan¹⁰ reported a significant increase in serial AoP measurements in cases of vaginal delivery, while progression remained minimal in deliveries resulting in cesarean section. ROC analyses identified a threshold value of >132.5° for AoP as a strong predictor. Chan et al.¹¹ demonstrated that AoP in prolonged second stage is highly valuable in predicting the mode of delivery. AoP was found to be wider during both rest and pushing in women who had successful instrumental deliveries; values above 138.7° (rest) and 160.9° (pushing) provided high sensitivity and specificity. Hans et al.¹² reported that an AoP ≥120° when cervical

dilation was ≥ 6 cm provided high accuracy in predicting vaginal delivery, and that an increase of $\geq 10^\circ$ in AoP (delta AoP) was the strongest predictor of mode of delivery. Katzir et al.¹³ In nulliparous women with a long second stage, they reported that vaginal delivery rates significantly increased when AoP was $\geq 127^\circ$, and that each increase in AoP raised the likelihood of vaginal delivery. Nassr et al.¹⁴ emphasized in a systematic review and meta-analysis that AoP at the beginning of the second stage has the highest accuracy in the 120-140° range, with specificity further increasing in the 141-153° range. Jung et al.¹⁵ In a recent review, it was stated that AoP is reliable and reproducible in assessing fetal head descent, and that wider AoP values show a positive correlation with successful vaginal delivery and shorter delivery time. It was stated that thresholds of 110-115° are critical in prolonged first stage, while the 141-153° range provides high specificity in the second stage.

In our study, AoP was found to be consistently higher and its rate of increase more pronounced from the active phase onwards in women delivering vaginally. This result is largely consistent with the findings reported in the literature by Bulut and Ceyhan,¹⁰ Chan et al.,¹¹ Hans et al.,¹² Katzir et al.,¹³ Nassr et al.,¹⁴ and Jung and Lee;¹⁵ in particular, the defined threshold values (~110-115° in the early phase, 141-153° in the second stage) correspond to our data.

In our study, it was found that HPD values were consistently lower in women who gave birth vaginally during active labor. It was observed that HPD could be a helpful parameter in determining the mode of delivery in both nulliparous and multiparous women.

Kartal Gölcük et al.⁸ reported that HPD differed according to mode of delivery in serial measurements, with progression being significantly slower in the cesarean group. HPD >38 mm has been reported to be a strong predictor for cesarean delivery. Carvalho Neto et al.¹⁶ showed that HPD was higher in the cesarean group, particularly in the second stage. They reported that cutoff values of 43 mm for HPD in the second stage provided high sensitivity and specificity. Tse et al.¹⁷ reported in a large prospective study that HPD decreased rapidly and significantly in women who delivered vaginally.

Our findings are consistent with the literature. Like Kartal Gölcük et al.,⁸ we also found that HPD was lower in the vaginal delivery group and higher in the cesarean section group. As reported in the studies by Carvalho Neto et al.¹⁶ and Tse et al.,¹⁷ the decrease in HPD as labor progressed was also confirmed in our data.

In our study, MLA values were found to be significantly lower in women who had vaginal deliveries compared to the cesarean section group. However, the fact that the increase in MLA is associated with an unfavorable course of delivery suggests that this parameter may provide clinical benefits in intrapartum management.

Sainz et al.¹⁸ demonstrated in their prospective study conducted prior to vacuum-assisted delivery that MLA was strongly associated with the difficulty and failure of the procedure. Mean MLA values were found to be 35° in the easy vacuum group, 55° in the difficult vacuum group, and

76° in the failed vacuum group, emphasizing that MLA values $>45^\circ$ are a negative predictor associated with failed attempts. According to findings from Dall'Asta et al.,¹⁹ MLA assessed intrapartum was predictive of spontaneous vaginal birth in nulliparous women with extended second-stage labor. A narrower MLA measurement was strongly associated with higher rates of vaginal delivery. A comprehensive systematic review and meta-analysis published by Skinner et al.²⁰ in 2023 showed that the clinical value of MLA becomes apparent, especially when measured during pushing. Accordingly, it was reported that MLA measured at rest provided limited accuracy in predicting complicated vaginal delivery (AUC ≈ 0.64), but MLA measured during pushing offered much stronger predictive power (AUC ≈ 0.90). These findings highlight the importance of MLA, particularly in decision-making processes during the second stage.

In our study, the wider distribution of MLA in the cesarean section group is consistent with findings reported in the literature and supports the view that this parameter is a valuable indicator for predicting the mode of delivery.

Limitations

This study has some limitations. The small number of women in the cesarean group may have reduced statistical power and limits the generalizability of the findings. As the study was conducted in a single tertiary center, selection bias cannot be excluded. Ultrasound assessments were performed by one operator, and interobserver variability was not evaluated. Moreover, the observational design precludes causal inference, and neonatal outcomes were not included in the analysis.

CONCLUSION

Serial intrapartum ultrasound measurements of AoP, HPD and MLA provide objective predictors of delivery mode during active labor. Higher AoP values were associated with vaginal delivery, whereas increased HPD and MLA correlated with cesarean section. Notably, an early intrapartum AoP value of $<100^\circ$ was linked to a greater likelihood of cesarean delivery, underscoring its potential as an early predictive marker. Routine incorporation of these sonographic parameters may improve the quality of intrapartum clinical decisions and facilitate more efficient management of labor.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Adana City Hospital Scientific Researches Ethics Committee (Date: 25.07.2024, Decision No: 76).

Informed Consent

All patients signed and free and informed consent form.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Uwamariya J, Abayisenga G, Ndizeye N, et al. Effect of active follow-up of women with previous cesarean delivery on uptake of timely safe obstetric and surgical care: comparison between pre-intervention and intervention cohorts in Rwanda. *PLoS One*. 2025;20(6):e0325884. doi:10.1371/journal.pone.0325884
- Pinton A, Severac F, Meyer N, et al. A comparison of vaginal ultrasound and digital examination in predicting preterm delivery in women with threatened preterm labor: a cohort study. *Acta Obstet Gynecol Scand*. 2017;96(4):447-453. doi:10.1111/aogs.13071
- Haumonte JB, Blanc J, Castel P, et al. Uncertain fetal head engagement: a prospective randomized controlled trial comparing digital exam with angle of progression. *Am J Obstet Gynecol*. 2022;227(4):625.e1-625.e8. doi:10.1016/j.ajog.2022.04.018
- Nebel S, Vardon D, Dreyfus M, Pizzoferrato AC. 2D-transperineal ultrasound in delivery room: contribution in assessing labor progress, predicting outcome of labor and recognizing obstetric anal sphincter injuries (OASIS). A systematic review. *J Gynecol Obstet Hum Reprod*. 2025;54(3):102910. doi:10.1016/j.jogoh.2025.102910
- Usman S, Hanidu A, Kovalenko M, Hassan WA, Lees C. The sonopartogram. *Am J Obstet Gynecol*. 2023;228(5s):S997-s1016. doi:10.1016/j.ajog.2022.06.027
- Conversano F, Di Trani MG, Morello R, et al. Automated approach for enhancing fetal head station assessment in labor with transperineal ultrasound. *Ultrasound Med Biol*. 2025;51(1):180-187. doi:10.1016/j.ultrasmedbio.2024.09.022
- Wong KW, Sultan AH, Andrews V, Allen-Coward H, Thakar R. Does the fetal head station and position affect the pelvic floor muscles in labour? A prospective study using 3 dimensional transperineal ultrasound. *Eur J Obstet Gynecol Reprod Biol*. 2025;313:114577. doi:10.1016/j.ejogrb.2025.114577
- Kartal Gölcük E, Dincgez B, Ozgen G, Ustunyurt E. The predictive role of serial transperineal sonography during the first stage of labor for cesarean section. *J Perinat Med*. 2025;53(2):132-139. doi:10.1515/jpm-2024-0450
- Ghi T, Eggebo T, Lees C, et al. ISUOG Practice Guidelines: intrapartum ultrasound. *Ultrasound Obstet Gynecol*. 2018;52(1):128-139.
- Bulut AN, Ceyhan V. Evaluation of the labour process with serial transperineal ultrasonography and prediction of the type of birth. *J Perinat Med*. 2020;49(1):36-42. doi:10.1515/jpm-2020-0323
- Chan VYT, Lau WL, So MKP, Leung WC. Measuring angle of progression by transperineal ultrasonography to predict successful instrumental and cesarean deliveries during prolonged second stage of labor. *Int J Gynaecol Obstet*. 2019;144(2):192-198. doi:10.1002/ijgo.12712
- Hans R, Reddy D, Shetty J. Serial intrapartum ultrasound to predict vaginal delivery using angle of progression and head- progression distance in term nulliparous women. *Eur J Obstet Gynecol Reprod Biol*. 2025;305:125-131. doi:10.1016/j.ejogrb.2024.12.014
- Katzir T, Brezinov Y, Khairish E, Hadad S, Vaisbuch E, Levy R. Intrapartum ultrasound use in clinical practice as a predictor of delivery mode during prolonged second stage of labor. *Arch Gynecol Obstet*. 2023;307(3):763-770. doi:10.1007/s00404-022-06469-5
- Nassr AA, Berghella V, Hessami K, et al. Intrapartum ultrasound measurement of angle of progression at the onset of the second stage of labor for prediction of spontaneous vaginal delivery in term singleton pregnancies: a systematic review and meta-analysis. *Am J Obstet Gynecol*. 2022;226(2):205-214.e2. doi:10.1016/j.ajog.2021.07.031
- Jung JE, Lee YJ. Intrapartum transperineal ultrasound: angle of progression to evaluate and predict the mode of delivery and labor progression. *Obstet Gynecol Sci*. 2024;67(1):1-16. doi:10.5468/ogs.23141
- Carvalho Neto RH, Viana Junior AB, Moron AF, Araujo Júnior E, Carvalho FHC, Feitosa HN. Assessment of the angle of progression and distance perineum-head in the prediction of type of delivery and duration of labor using intrapartum ultrasonography. *J Matern Fetal Neonatal Med*. 2021;34(14):2340-2348. doi:10.1080/14767058.2019.1666818
- Tse WT, Chaemsaitong P, Chan WWY, et al. Labor progress determined by ultrasound is different in women requiring cesarean delivery from those who experience a vaginal delivery following induction of labor. *Am J Obstet Gynecol*. 2019;221(4):335.e1-335.e18. doi:10.1016/j.ajog.2019.05.040
- Sainz JA, Borrero C, Aquise A, Serrano R, Gutiérrez L, Fernández-Palacín A. Utility of intrapartum transperineal ultrasound to predict cases of failure in vacuum extraction attempt and need of cesarean section to complete delivery. *J Matern Fetal Neonatal Med*. 2016;29(8):1348-1352. doi:10.3109/14767058.2015.1048680
- Dall'Asta A, Angeli L, Masturzo B, et al. Prediction of spontaneous vaginal delivery in nulliparous women with a prolonged second stage of labor: the value of intrapartum ultrasound. *Am J Obstet Gynecol*. 2019;221(6):642.e1-642.e13. doi:10.1016/j.ajog.2019.09.045
- Skinner SM, Giles-Clark HJ, Higgins C, Mol BW, Rolnik DL. Prognostic accuracy of ultrasound measures of fetal head descent to predict outcome of operative vaginal birth: a comparative systematic review and meta-analysis. *Am J Obstet Gynecol*. 2023;229(1):10-22.e10. doi:10.1016/j.ajog.2022.11.1294

Bilingual evaluation of YouTube™ videos on teething symptoms and remedies

 Deniz Demir¹,  Hülya Çerçi Akçay²

¹Dentistry Student, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Türkiye

²Department of Pedodontics, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Türkiye

Cite this article as: Demir D, Çerçi Akçay H. Bilingual evaluation of YouTube™ videos on teething symptoms and remedies. *J Health Sci Med.* 2025;8(6):1082-1088.

Received: 17.09.2025

Accepted: 09.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to evaluate the reliability, scientific accuracy, and educational value of YouTube™ videos related to teething symptoms and remedies for parents. Given the widespread use of digital platforms for health information, the study focused on identifying content quality differences between professional and non-professional sources.

Methods: A systematic search was performed on the YouTube™ platform (www.youtube.com) on July 25, 2025, using a computer in Türkiye. Both English (“teething symptoms,” “teething treatment”) and Turkish (“diş çıkarma semptomları,” “diş çıkarma tedavileri”) keywords were used. The first 100 videos retrieved for each search were screened, and duplicates, promotional content, and non-relevant materials were excluded. Eligible videos were assessed using the Global Quality Scale (GQS), DISCERN instrument, JAMA benchmark criteria, and the Video Information and Quality Index (VIQI). Metadata including video length, views, likes, dislikes, and uploader category was also collected. Statistical analyses included descriptive statistics, Mann-Whitney U tests, and Spearman’s correlation coefficients ($p < 0.05$).

Results: A total of 100 YouTube™ videos were analyzed. Mean scores were low across all instruments; GQS 2.6, DISCERN 11.7, JAMA 1.8, and VIQI 12.5. Videos from academic institutions achieved higher quality scores than non-professional sources. Likes showed significant positive correlations with GQS ($p = 0.315$) and DISCERN ($p = 0.360$, $p < 0.01$), while view counts were not associated with quality. Strong intercorrelations were observed among all quality tools ($p > 0.45$, $p < 0.01$).

Conclusion: YouTube™ provides a readily accessible platform for information on teething symptoms and remedies; however, the majority of videos lack scientific rigor, comprehensive explanations, and evidence-based recommendations. Content produced by healthcare professionals and academic institutions demonstrated superior quality, underscoring the need for their active involvement in creating reliable digital health resources.

Keywords: Educational technology, health communication, health knowledge, internet, pain management, teething, Youtube™

INTRODUCTION

Teething is a common developmental milestone during infancy, typically occurring between the ages of 4 to 36 months.^{1,2} This process involves the eruption of primary teeth, a natural yet challenging phase for both infants and their caregivers. As primary teeth emerge, various local and systemic symptoms may be observed, such as general irritability, sleep disturbances, crying, nasal discharge, flushed cheeks, fever, diarrhea, loss of appetite, hypersalivation, ear rubbing, and gingival inflammation on the erupting teeth, in this period of time when primary teeth erupt.³⁻⁵ While these symptoms are often viewed as normal during teething, they can be distressing for parents, leading them to seek ways to alleviate the discomfort their child experiences. Research has shown, however, that there is no significant causal relationship between teething and the occurrence of systemic symptoms like fever, diarrhea, or infections.^{3,6,7} Despite this, the symptoms of teething, especially local ones like irritability

and discomfort, can still cause concern, which often drives parents to look for relief options.

To alleviate the discomfort associated with teething, several pharmacological and non-pharmacological interventions have been recommended. Pharmacological treatments, such as analgesics and antipyretics, may be effective in managing pain and fever, while non-pharmacological strategies are also considered important adjuncts. These include the use of teething rings, chilled fruits and vegetables, and massaging the gums with clean fingers, cold spoons, or wet gauze. Additionally, pacifiers, topical anesthetic agents, and even alternative holistic medicine are frequently used by parents seeking relief for their children during this challenging period.⁶ Despite the availability of various remedies, scientific evidence supporting the efficacy of many of these treatments remains limited, and parents are often left to navigate a complex landscape of conflicting advice.

Corresponding Author: Hülya Çerçi Akçay, hulya.akcay@kocaelisağlik.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

Parents typically seek guidance on teething symptoms from healthcare providers such as pediatric dentists, pediatricians, family physicians, and other medical professionals. However, in recent years, there has been a marked increase in the use of online resources for obtaining health information related to infant care. The teething period requires careful monitoring, as it presents challenges for both infants and their caregivers.⁸⁻¹⁰ This shift toward digital platforms for health advice has introduced a new dynamic, with YouTube™ emerging as a widely-used and influential resource. YouTube™, the world's second most frequently used search engine and one of the most widely utilized social media platforms, has become a go-to source for parents seeking information on teething and other aspects of child health.¹¹ The ease of access to vast amounts of content on YouTube™ makes it an attractive tool for parents, but the lack of expert oversight means that the quality and reliability of the information available can vary significantly. As such, it is crucial to critically evaluate the accuracy and credibility of the information being disseminated on such platforms.

This growing reliance on video-based platforms for health-related information underscores the importance of understanding the quality of the content available, particularly when it comes to critical topics like teething, which requires careful management. With parents increasingly turning to YouTube™ for advice, the question arises as to whether the platform serves as a reliable source of health information. In order to provide parents with the most accurate and useful content, it is essential to assess whether YouTube™ videos provide evidence-based guidance and adhere to medical and scientific standards.

This study aims to evaluate the quality, accuracy, and scientific credibility of YouTube™ videos as an informational resource for parents seeking guidance on teething.

METHODS

Ethical Statement

This study did not involve human participants or identifiable personal data. Only publicly available YouTube™ content was analyzed. Therefore, ethical approval was not required, in accordance with the principles of the Declaration of Helsinki.

Video Selection and Data Collection

A systematic search was conducted on the YouTube™ platform (www.youtube.com) on July 25, 2025, using a computer located in Türkiye. The purpose was to identify publicly accessible videos providing information on teething symptoms and potential treatments for children. To reflect the bilingual search behavior of parents, both English and Turkish videos were evaluated.

The search terms used were “teething symptoms” and “teething treatment” for English content, and “diş çıkarma semptomları” and “diş çıkarma tedavileri” for Turkish content. These keywords were selected based on both existing literature and common user search patterns to capture the terminology most likely employed by the general public.

Prior to the search, browser history and cookies were cleared to minimize personalization bias, and the default “Sort by

relevance” filter on YouTube™ was applied. Because users predominantly engage with only the first three pages of search results, the first 100 videos retrieved for each keyword were included. This approach is supported by Desai et al.,¹² who reported that approximately 95% of viewers restrict their searches to this range.

For each video, metadata-including duration, upload date, view count, likes, dislikes, number of comments, and uploader category-was systematically recorded. Videos were classified into four source groups: (1) individual healthcare professionals, (2) universities or academic institutions, (3) health information platforms, and (4) other individual users.

Eligibility Criteria

To ensure methodological consistency and reliability, strict inclusion and exclusion criteria were applied. Videos were included if they met the following conditions: (1) English or Turkish language content, (2) minimum resolution of 240 p, (3) duration less than 30 minutes, and (4) a primary focus on teething symptoms and/or treatments. Videos were required to contain explanatory content delivered through narration, subtitles, or visual demonstration to qualify as educational material.

Videos were excluded if they were promotional in nature, created for entertainment purposes, of insufficient visual quality, or unrelated to teething. Duplicate videos and YouTube “shorts” were also excluded because of their limited duration and restricted educational content.

The inclusion and exclusion process was conducted by two independent researchers. To assess inter-rater reliability, both Cohen's kappa (κ) statistic and Intraclass Correlation Coefficient (ICC) were applied depending on the type of scale. Specifically, Weighted Kappa was used for ordinal scales such as the Global Quality Scale (GQS) and DISCERN score, while ICC (two-way mixed effects, absolute agreement model) was employed for continuous variables including JAMA, VIQI, and the total score. The overall inter-rater agreement was substantial, with Cohen's kappa calculated as $\kappa=0.81$ for categorical ratings. Disagreements were resolved by consensus, with a third reviewer consulted if necessary.

Statistical Analysis

All data analyses were performed using Python (version 3.11). As the data did not conform to a normal distribution, non-parametric tests were employed. Specifically, Spearman's rank-order correlation was used to assess associations between viewer engagement metrics (e.g., views, likes, dislikes, interaction index) and quality assessment scores (e.g., GQS, DISCERN, JAMA, VIQI). Correlation coefficients (ρ) and associated p-values were reported, with statistical significance set at $p<0.05$.

Descriptive statistics, including means, standard deviations, medians, and interquartile ranges (IQR), were calculated to summarize the distribution of variables. To illustrate variation across uploader types, boxplot visualizations were generated for key metrics such as the number of likes and the Interaction Index.

Data processing and visualization were conducted using Python libraries, including pandas for data handling, scipy.stats for statistical computation, and matplotlib and seaborn for plotting.

RESULTS

Descriptive Statistics

A total of 100 YouTube™ videos were included in the final analysis, of which 56 were in Turkish and 44 were in English. These videos were systematically assessed for content quality using validated evaluation tools, including the GQS, DISCERN instrument, JAMA Benchmark Criteria, and the Video Information and Quality Index (VIQI). Additionally, viewer engagement metrics-such as total number of views, likes, dislikes, and Interaction Index-were recorded and analyzed to evaluate the popularity and audience response toward the content (Table 1).

Table 1. Descriptive statistics of Turkish and English YouTube™ videos on teething				
Parameter	English (mean±SD)	English (median, IQR)	Turkish (mean±SD)	Turkish (median, IQR)
Total views	112.085±194.931	18.465 (157.988)	44.29±111.562	3.410 (18.390)
Likes	811±1.597	145 (963)	797±2.787	20 (94)
Dislikes	0.0±0.0	0 (0)	0.0±0.0	0 (0)
Engagement Index	1.04±1.35	0.66 (0.84)	1.47±4.00	0.40 (1.33)
SD: Standard deviation, IQR: Interquartile range. The table summarizes descriptive statistics of Turkish and English YouTube™ videos related to teething, including total number of views, likes, dislikes, and Engagement Index.				

Quality Scores

The mean DISCERN score of the evaluated videos was 11.7 (SD=4.2), indicating a generally low-to-moderate level of reliability and scientific accuracy. Similarly, the mean GQS was 2.6 (SD=1.1), suggesting that the videos tended to offer limited comprehensiveness and educational value. The average JAMA score was 1.8 (SD=0.6), highlighting the frequent absence of essential authorship, attribution, disclosure, and currency criteria. VIQI scores were also modest, with a mean of 12.5 (SD=3.6), reflecting generally poor structural coherence and audiovisual quality (Table 2).

Table 2. Quality assessment scores for both Turkish and English YouTube™ videos on teething symptoms and remedies		
Parameter	Mean±SD	Median (IQR)
Global Quality Scale	2.6±1.1	2 (1)
DISCERN score	11.7±4.2	11 (5)
JAMA score	1.8±0.6	2 (1)
Video Information and Quality Index	12.5±3.6	13 (6)
SD: Standard deviation; IQR: Interquartile range. The table presents descriptive statistics of Global Quality Scale, DISCERN, JAMA, and Video Information and Quality Index scores.		

When categorized by uploader type, videos uploaded by academic institutions had the highest mean scores across all quality indicators, particularly for GQS and DISCERN, suggesting that institutional content is more likely to meet established educational and scientific standards. In contrast,

videos uploaded by individual users and health-related websites demonstrated wide variability in quality, with several scoring poorly on all evaluation scales.

Correlation Analysis

Spearman correlation coefficients were computed to explore the associations between viewer engagement metrics (e.g., views, likes, dislikes, interaction index) and standardized video quality scores (DISCERN, GQS, JAMA, VIQI). The correlation matrix revealed several statistically significant relationships across these variables.

The total number of views exhibited a very strong positive correlation with the number of likes ($p=0.883$, $p<0.01$), indicating a strong alignment between video visibility and user engagement in the form of likes. However, no significant correlations were observed between view count and any of the four quality evaluation instruments: DISCERN ($p=0.07$), GQS ($p=0.09$), JAMA ($p=0.16$), and VIQI ($p=0.13$), suggesting limited association between view-based popularity and standardized content quality scores.

The number of likes showed statistically significant positive correlations with all four quality measures. Specifically, the correlation with DISCERN was $p=0.360$ ($p<0.01$), with GQS $p=0.315$ ($p<0.01$), with JAMA $p=0.291$ ($p<0.01$), and with VIQI $p=0.206$ ($p<0.05$). These correlations suggest that videos receiving higher numbers of likes tended to score better in terms of reliability, global quality, structural rigor, and visual information quality.

Interestingly, the number of dislikes also showed statistically significant positive associations with quality indicators. Dislike counts were correlated with DISCERN ($p=0.283$, $p<0.01$), GQS ($p=0.208$, $p<0.05$), and JAMA ($p=0.276$, $p<0.01$). No significant correlation was observed between dislikes and VIQI ($p=0.187$, $p=0.07$), indicating a marginal association.

The Interaction Index, which aggregates likes, comments, and views, demonstrated statistically significant moderate correlations with DISCERN ($p=0.346$, $p<0.01$) and GQS ($p=0.312$, $p<0.01$), but did not exhibit statistically significant correlations with JAMA ($p=0.199$, $p=0.06$) or VIQI ($p=0.175$, $p=0.13$). Additionally, the interaction index did not significantly correlate with the total view count ($p=0.63$), suggesting that aggregate engagement may not be directly influenced by raw visibility metrics.

Among the internal quality assessment instruments, strong and statistically significant correlations were observed, indicating high inter-rater consistency and overlapping domains of evaluation. Specifically:

- GQS was significantly correlated with DISCERN ($p=0.649$, $p<0.01$), JAMA ($p=0.544$, $p<0.01$), and VIQI ($p=0.451$, $p<0.01$).
- DISCERN showed strong positive correlations with JAMA ($p=0.591$, $p<0.01$) and VIQI ($p=0.590$, $p<0.01$).
- JAMA was also positively correlated with VIQI ($p=0.457$, $p<0.01$).

Collectively, these results indicate that the standardized instruments used to evaluate structural and educational

quality were strongly interrelated, producing internally consistent scores across different dimensions of content evaluation (Table 3 and Figure 1, 2).

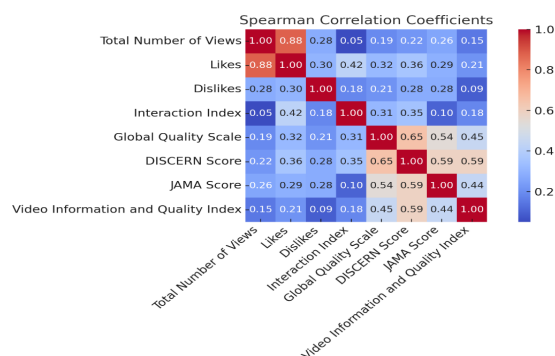


Figure 1. Spearman correlation coefficients among quantitative YouTube video metrics and quality evaluation scores (GQS, DISCERN, JAMA, VIQI). Positive correlations are indicated in red and negative correlations in blue. Stronger color intensity corresponds to higher correlation values

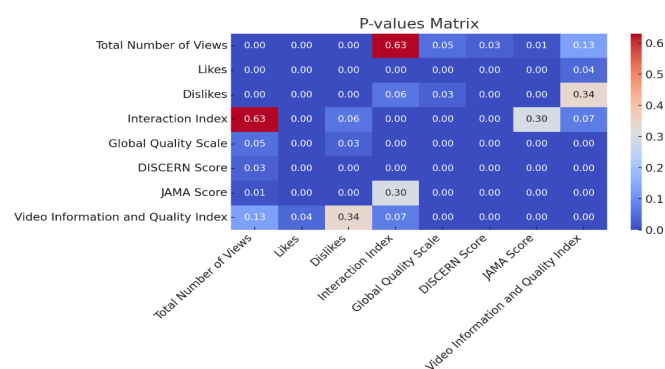


Figure 2. p-values matrix corresponding to the Spearman correlation between YouTube video metrics and quality assessment scores. Statistically significant relationships ($p < 0.05$) are represented in dark blue, while non-significant correlations are shown in lighter shades

Boxplot Analysis by Uploader Type

Boxplot visualizations were generated to compare the distribution of like counts and Interaction Index values across different video uploader types, including academic institutions, dentists, individual users, and health-related websites.

The distribution of like counts showed that videos uploaded by academic institutions had a higher median number of likes compared to those from other uploader types. The interquartile

range for academic sources was narrower, indicating more consistent approval metrics, whereas individual users and dentists displayed broader variability and outliers in like counts.

For the Interaction Index, which aggregates likes, comments, and views, the boxplot indicated greater variability among individual users and dentists. Some videos from these groups exhibited higher interaction scores, as reflected by wider interquartile ranges and the presence of extreme values. In contrast, videos from academic sources showed lower dispersion and fewer outliers in Interaction Index values.

These distributional differences between uploader categories were visualized to facilitate further interpretation of engagement dynamics across source types (Figure 3).

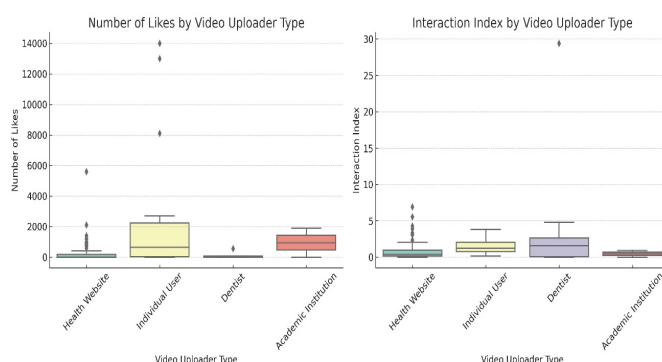


Figure 3. Boxplots comparing the number of likes and interaction index across different video uploader types

Inter-observer Reliability

Interobserver agreement was evaluated to ensure rating consistency. As presented in Table 4, Weighted Kappa values showed moderate agreement for the Global Quality Scale (GQS, $\kappa = 0.599$, 95% CI: 0.431-0.768, $p < 0.001$) and substantial agreement for DISCERN ($\kappa = 0.783$, 95% CI: 0.678-0.888, $p < 0.001$). The intraclass correlation coefficients demonstrated excellent inter-rater reliability for the JAMA benchmark criteria (ICC=0.921, 95% CI: 0.864-0.955, $p < 0.001$) and for the Video Information and Quality Index (VIQI, ICC=0.915, 95% CI: 0.853-0.951, $p < 0.001$). The composite total score, defined as the sum of DISCERN, JAMA, and VIQI scores, achieved the highest level of agreement with an ICC of 0.965 (95% CI: 0.938-0.980, $p < 0.001$), reflecting strong alignment between evaluators across all scales.¹³

Table 3. Spearman correlation matrix with significance levels between YouTube engagement metrics and quality assessment tools (Global Quality Scale, DISCERN, JAMA, Video Information and Quality Index). Correlation is significant at the 0.01 level (2-tailed) is indicated by "*" and at the 0.05 level (2-tailed) by "*".

	Total number of views	Likes	Dislikes	Engagement Index	Global Quality Scale	DISCERN score	JAMA score	Video Information and Quality Index
Total number of views	1.000**	0.883**	0.278**	0.047	0.193*	0.216*	0.257**	0.150
Likes	0.883**	1.000**	0.305**	0.420**	0.315**	0.360**	0.291**	0.206*
Dislikes	0.278**	0.305**	1.000**	0.182	0.208*	0.283**	0.276**	0.094
Engagement Index	0.047	0.420**	0.182	1.000**	0.312**	0.346**	0.102	0.179
Global Quality Scale	0.193*	0.315**	0.208*	0.312**	1.000**	0.649**	0.544**	0.451**
DISCERN score	0.216*	0.360**	0.283**	0.346**	0.649**	1.000**	0.585**	0.590**
JAMA score	0.257**	0.291**	0.276**	0.102	0.544**	0.585**	1.000**	0.444**
Video Information and Quality Index	0.150	0.206*	0.094	0.179	0.451**	0.590**	0.444**	1.000**

Table 4. Interobserver agreement levels for Video Evaluation Scales using weighted Kappa and ICC

Scale	ICC/weighted Kappa	95% CI	p-value
Global Quality Scale (weighted Kappa)	0.668	0.431-0.768	<0.001
DISCERN (weighted Kappa)	0.908	0.678-0.888	<0.001
JAMA (ICC)	0.921	0.864-0.955	<0.001
Video Information and Quality Index (ICC)	0.849	0.853-0.951	<0.001
Total score (ICC)	0.965	0.938-0.980	<0.001

Note: Weighted Kappa was used for the Global Quality Scale and DISCERN score, while ICC (two-way mixed effects, absolute agreement) was applied for the JAMA, VIQI, and total score. ICC: Intraclass correlation coefficient, CI: Confidence interval

DISCUSSION

The widespread use of digital platforms has significantly influenced how parents access health-related information about their infants. In parallel with the ongoing technological advancements and the digitalization of healthcare services, platforms such as YouTube™ have become increasingly popular sources of medical content.¹⁴ This shift has created new opportunities for disseminating health information but also raised concerns regarding the accuracy, quality, and reliability of user-generated content.¹⁵

Previous studies in pediatric dentistry have evaluated YouTube™ videos on topics such as oral hygiene, early childhood caries, and fluoride use.^{3,16} In addition, some research has analyzed teething-related content, but these investigations were limited to English-language videos only.^{5,17}

Unlike those studies, the present study incorporated both Turkish and English videos, thereby offering a broader bilingual perspective. While this approach is novel in including content from two different languages, it should be emphasized that no direct comparative analysis between English- and Turkish-language videos was performed. Instead, only basic descriptive statistics such as the number of videos, mean views, likes, dislikes, and video duration were presented separately for each language group. To our knowledge, this is the first study in pediatric dentistry to extend video evaluations to both Turkish and English content, thereby broadening the scope of previous single-language analyses.

When categorized by uploader type, videos produced by academic institutions consistently achieved the highest mean scores across multiple quality assessment metrics, particularly in the GQS and the DISCERN instrument. This finding highlights the superiority of academic sources in adhering to scientific methodology and pedagogical standards, as they are generally developed with a more rigorous and evidence-based approach. In contrast, videos uploaded by individual users and health-focused websites demonstrated substantial heterogeneity in informational integrity and accuracy, with a significant portion performing poorly across evaluation criteria. Similar trends have been observed in previous studies evaluating YouTube™ videos on fluoride therapy, early childhood caries, teeth whitening, and pediatric surgery, all of which emphasized that professionally

generated content is more reliable and educational than layperson-generated videos.¹⁸⁻²¹ This variability reinforces the risk of misinformation in digital health communication and underscores the need for regulatory frameworks and systematic quality control protocols to safeguard information quality online.

The present study further revealed considerable methodological and educational limitations in YouTube™ videos addressing teething symptoms and remedies. Low mean scores across DISCERN (11.7), GQS (2.6), JAMA (1.8), and VIQI (12.5) collectively indicate that most analyzed videos failed to meet recognized standards of evidence-based health communication. Deficiencies were particularly evident in authorship transparency, source attribution, structural coherence, and audiovisual quality. Comparable shortcomings have also been reported in previous analyses of pediatric dentistry-related content, including oral hygiene and oral habits, where lack of referencing and poor audiovisual standards compromised reliability.^{22,23}

Another noteworthy finding was the absence of statistically significant correlations between total view count and content quality, suggesting that video visibility is algorithm-driven rather than associated with scientific reliability. Conversely, likes were moderately and positively correlated with all four quality indices, indicating that audience approval may partially reflect perceived educational value. Surprisingly, dislikes also showed positive correlations with DISCERN, GQS, and JAMA scores, although not with VIQI, suggesting that negative user feedback is not always directed at low-quality information but may instead reflect subjective disagreements or stylistic preferences.

Overall, these findings underscore a critical gap between user engagement signals and actual content quality. While view counts strongly correlate with popularity, they do not reliably indicate scientific rigor. By contrast, likes and dislikes demonstrated moderate associations with standardized quality indicators, reflecting a more nuanced relationship between user feedback and credibility.

From a public health perspective, these findings are concerning. Teething, although physiologically benign, often provokes significant parental anxiety. When parents rely on low-quality digital resources, they risk exposure to misinformation, adoption of inappropriate home remedies, or unnecessary anxiety-driven medical consultations.^{5,17} Moreover, the absence of alignment between video popularity and scientific accuracy highlights the algorithmic vulnerability of digital platforms, where engagement cues rather than reliability drive visibility.²⁴

Given these risks, academic institutions and professional organizations must take a more proactive role in digital content creation. Developing evidence-based, accessible, and linguistically diverse resources is essential. Practical strategies may include implementing quality verification badges, leveraging algorithms to prioritize peer-reviewed content, and establishing collaborations with platform providers to enhance visibility of institutionally verified materials.²⁵

Further analysis of user engagement revealed that the Interaction Index-a composite of likes, comments, and views-showed significant moderate correlations with DISCERN and GQS, but no significant associations with JAMA or VIQI. Importantly, no relationship was found between interaction index and total view count, underscoring that active engagement is not directly linked to raw viewership. Strong and significant correlations were observed among all four quality assessment tools, indicating that these instruments consistently capture overlapping methodological and educational quality dimensions.

Future research should extend beyond YouTube™ to encompass platforms such as TikTok, Instagram, and Facebook, where health-related content is increasingly consumed.^{26,27} Longitudinal studies are also needed to evaluate changes in video quality and user engagement over time, particularly in response to evolving platform algorithms and public health interventions.²⁸ Additionally, intervention-based research is warranted to test the effectiveness of strategies such as content verification systems, digital literacy campaigns, or educational labeling in improving the reliability and educational value of health-related digital content.²⁹

Limitations

This study has several limitations. First, it was restricted to YouTube™, excluding other popular platforms such as TikTok and Instagram. Second, the cross-sectional design provides only a snapshot in time; video quality and availability may change. Third, while both Turkish- and English-language videos were included, no direct statistical comparison was performed; only descriptive information was presented. Finally, the analysis was limited to publicly available videos, which may affect generalizability.

Future research should therefore adopt longitudinal and multi-platform approaches, with more detailed cross-linguistic analyses to strengthen the robustness of findings.

CONCLUSION

YouTube™ videos on teething frequently lack accuracy and educational value, creating risks of parental misinformation. Pediatric dental professionals should actively contribute evidence-based content. Equally important, digital platforms must adopt policy-level quality control strategies-such as verification systems and algorithmic prioritization of reliable sources-to ensure that trustworthy health information reaches parents. These combined efforts are essential to strengthen parental health literacy and support safe, evidence-informed childcare practices.

ETHICAL DECLARATIONS

Ethics Committee Approval

Publicly available YouTube™ content was analyzed and no ethical approval was required.

Informed Consent

Because the study has no study with human and human participants, no written informed consent form was obtained.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Acknowledgment

This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) under the 2209-A Research Project Program, project number 1919B012405673.

REFERENCES

1. Marks SC Jr. The basic and applied biology of tooth eruption. *Connect Tissue Res.* 1995;32(1-4):149-157. doi:10.3109/03008209509013718
2. Tsang AK, Annetta KL. Teething, teething pain and teething remedies. *Int Dent South Africa.* 2010;12(5):48-61.
3. Markman L. Teething: facts and fiction. *Pediatr Rev.* 2009;30(8):e59-e64. doi:10.1542/pir.30-8-e59
4. Macknin ML, Piedmonte M, Jacobs J, Skibinski C. Symptoms associated with infant teething: a prospective study. *Pediatrics.* 2000;105(4 Pt 1):747-752. doi:10.1542/peds.105.4.747
5. Güçyetmez Topal B, Tıraş M, Tanrikulu A. Assessment of YouTube videos as a parent information source for teething symptoms. *Cumhuriyet Dent J.* 2022;25(Suppl):114-118. doi:10.7126/cumudj.1032447
6. Ramos-Jorge J, Pordeus IA, Ramos-Jorge ML, Paiva SM. Prospective longitudinal study of signs and symptoms associated with primary tooth eruption. *Pediatrics.* 2011;128(3):471-476. doi:10.1542/peds.2010-2697
7. Wake M, Hesketh K, Lucas J. Teething and tooth eruption in infants: a cohort study. *Pediatrics.* 2000;106(6):1374-1379. doi:10.1542/peds.106.6.1374
8. Wake M, Hesketh K, Allen M. Parent beliefs about infant teething: a survey of Australian parents. *J Paediatr Child Health.* 1999;35(5):446-449. doi:10.1046/j.1440-1754.1999.355395.x
9. Kozuch M, Peacock E, D'Auria JP. Infant teething information on the world wide web: taking a byte out of the search. *J Pediatr Health Care.* 2015;29(1):38-45. doi:10.1016/j.pedhc.2014.06.006
10. Walsh AM, Hamilton K, White KM, Hyde MK. Use of online health information to manage children's health care: a prospective study investigating parental decisions. *BMC Health Serv Res.* 2015;15:131. doi:10.1186/s12913-015-0793-4
11. Atkinson NL, Saperstein SL, Pleis J. Using the internet for health-related activities: findings from a national probability sample. *J Med Internet Res.* 2009;11(1):e4. doi:10.2196/jmir.1035
12. Desai T, Shariff A, Dhingra V, Minhas D, Eure M, Kats M. Is content really king? An objective analysis of the public's response to medical videos on YouTube. *PLoS One.* 2013;8(12):e82469. doi:10.1371/journal.pone.0082469
13. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-174.
14. Boulous MN, Brewer AC, Karimkhani C, Buller DB, Dellavalle RP. Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J Public Health Inform.* 2014;5(3):229. doi:10.5210/ojphi.v5i3.4814
15. Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. *Health Informatics J.* 2015;21(3):173-194. doi:10.1177/1460458213512220

16. ElKarmi R, Hassona Y, Taimeh D, Scully C. YouTube as a source for parents' education on early childhood caries. *Int J Paediatr Dent*. 2017; 27(6):437-443. doi:10.1111/ipd.12277
17. Eroglu Cakmakoglu E, Bakir M. Evaluation of YouTube videos about teething. *Akd Med J*. 2024;10(3):426-433. doi:10.53394/akd.1310739
18. Eglil E, Altan Sallı G. YouTube as a source of information on fluoride therapy. *Fluoride*. 2020;53(2):292-301.
19. Simsek H, Buyuk SK, Cetinkaya E, Tural M, Koseoglu MS. How I whiten my teeth: YouTube as a patient information resource for teeth whitening. *BMC Oral Health*. 2020;20(1):1-6. doi:10.1186/s12903-020-01172-w
20. Bezner SK, Hodgman EI, Diesen DL, et al. Pediatric surgery on YouTube: is the truth out there? *J Pediatr Surg*. 2014;49(4):586-589. doi:10.1016/j.jpedsurg.2013.08.004
21. Duman C. YouTube quality as a source for parent education about the oral hygiene of children. *Int J Dent Hyg*. 2020;18(3):261-267. doi:10.1111/idh.12445
22. Simsek H, Buyuk SK, Cetinkaya E. YouTube as a source of information on oral habits. *J Indian Soc Pedod Prev Dent*. 2020;38(2):115-118. doi:10.4103/JISPPD.JISPPD_357_19
23. Sampson M, Cumber J, Li C, Pound CM, Fuller A, Harrison D. A systematic review of methods for studying consumer health YouTube videos, with implications for systematic reviews. *PeerJ*. 2013;1:e147. doi:10.7717/peerj.147
24. Gholami-Kordkheili F, Wild V, Strech D. The impact of social media on medical professionalism: a systematic qualitative review of challenges and opportunities. *J Med Internet Res*. 2013;15(8):e184. doi:10.2196/jmir.2708
25. Butler DP, Perry F, Shah Z, Leon-Villapalos J. The quality of video information on burn first aid available on YouTube. *Burns*. 2013;39(5):856-861. doi:10.1016/j.burns.2012.10.017
26. Basch CH, Hillyer GC, Jaime C. COVID-19 on TikTok: harnessing an emerging social media platform to convey important public health messages. *Int J Adolesc Med Health*. 2020;34(5):367-369. doi:10.1515/ijamh-2020-0111
27. Peterson EB, Chou WS, Rising C, Gaysynsky A. The role and impact of health literacy on peer-to-peer health communication. *Stud Health Technol Inform*. 2020;269:497-510. doi:10.3233/SHTI200058
28. Uzel İ, Ghabchi B, Akalın A, Eden E. YouTube as an information source in paediatric dentistry education: reliability and quality analysis. *PLoS One*. 2023;18(3):e0283300. doi:10.1371/journal.pone.0283300
29. Şahin TN, Özmen EE. Evaluation of online information on bruxism in children: a study on YouTube videos. *Med Records*. 2024;6(3):347-353. doi:10.37990/medr.1480011

Separate and combined effects of maternal obesity and preeclampsia on perinatal outcomes: a retrospective four-group analysis

 Erkan Sağlam

Department of Perinatology, Bursa City Hospital, Bursa, Türkiye

Cite this article as: Sağlam E. Separate and combined effects of maternal obesity and preeclampsia on perinatal outcomes: a retrospective four-group analysis. *J Health Sci Med.* 2025;8(6):1089-1098.

Received: 23.07.2025

Accepted: 10.10.2025

Published: 25.10.2025

ABSTRACT

Aims: Obesity and preeclampsia are common complications that can negatively affect maternal and perinatal outcomes. However, their combined impact and the modifying role of gestational weight gain (GWG) remain unclear.

Methods: This retrospective cohort included 400 women classified into four groups by first-trimester BMI (obesity ≥ 30 kg/m²) and the presence of preeclampsia (PE): PE(-) Ob(-), PE(+) Ob(-), PE(-) Ob(+), and PE(+) Ob(+). From 2,600 eligible records, we used stratified random sampling to select 100 women per group (no covariate matching) to ensure balanced precision. Demographic, obstetric, and neonatal outcomes were compared across groups. Risks of preterm birth (<37 and <34 weeks), low birth weight, neonatal intensive care unit (NICU) admission, and low Apgar scores were estimated with logistic regression adjusted for maternal age, smoking, in vitro fertilization, and gestational weight gain (GWG); a PE×Ob interaction term was tested. To examine GWG as a continuous exposure, we fitted restricted cubic spline models. Analyses were complete-case.

Results: The PE(+) Ob(+) group had the highest absolute event rates: preterm birth <37 weeks occurred in 38% compared with 5% in the PE(-) Ob(-) group; preterm birth <34 weeks in 12% vs 1%; low birth weight in 37% vs 4%; NICU admission in 25% vs 3%; 1-minute Apgar <7 in 22% vs 4%; and 5-minute Apgar <7 in 13% vs 2%. Adjusted analyses confirmed markedly elevated odds for these outcomes-preterm birth <37 weeks (aOR 12.17; 95% CI 4.36-33.97), preterm birth <34 weeks (aOR 18.44; 95% CI 2.24-151.86), low birth weight (aOR 19.61; 95% CI 6.22-61.81), NICU admission (aOR 12.41; 95% CI 3.45-44.62), and low Apgar scores at 1 and 5 minutes (all $p < 0.05$). The PE(+) Ob(-) group also showed increased risks for most of these outcomes except preterm birth <34 weeks and low 5-minute Apgar scores, whereas the PE(-) Ob(+) group did not demonstrate significant associations. Restricted cubic spline analysis suggested that inadequate gestational weight gain may further increase these risks in the PE(+) Ob(+) group.

Conclusion: Obesity alone was not associated with increased risk, whereas co-occurrence with preeclampsia was associated with substantially higher risks; no significant PE×Ob interaction was detected. Careful monitoring of gestational weight gain in such pregnancies is warranted.

Keywords: Preeclampsia, obesity, perinatal outcomes, gestational weight gain, preterm birth

INTRODUCTION

Obesity is an increasingly significant public health problem for women of reproductive age worldwide.¹ It adversely affects pregnancy outcomes for both the mother and the fetus.^{2,3} Obese women who are pregnant are more likely to get gestational diabetes, macrosomia, have a cesarean delivery, and have thromboembolic problems.⁴ Some studies have also found that obesity may be linked to bad outcomes during pregnancy and birth, like giving birth too early and having low Apgar scores.^{5,6}

Preeclampsia is a serious pregnancy-related high blood pressure condition that can harm both the mother and the fetus and lead to a lot of illness and death.⁷ Preeclampsia can cause bad consequences for babies and mothers, like premature birth, slow growth in the womb, not enough placenta, and the requirement for admission to the neonatal intensive

care unit (NICU).⁸ Obesity is an independent risk factor for preeclampsia, and obese women have a higher likelihood of developing the condition.⁹ Given this, the coexistence of obesity and preeclampsia is common, yet their combined impact warrants further investigation.

However, there is limited research comparing perinatal outcomes when obesity and preeclampsia coexist versus when each occurs independently. Moreover, the role of modifiable factors such as gestational weight gain (GWG) in this relationship remains unclear. Therefore, the aim of this study was to evaluate the impact of obesity and preeclampsia, individually and combined, on perinatal outcomes and to assess the potential modifying effect of gestational weight gain.

Corresponding Author: Erkan Sağlam, dresaglam@gmail.com



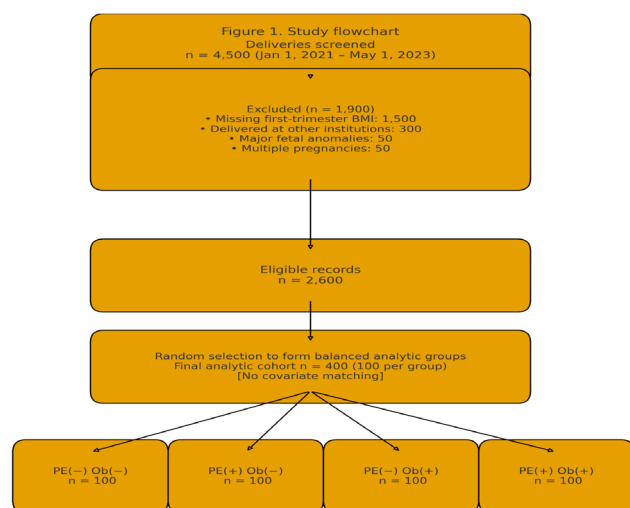
This work is licensed under a Creative Commons Attribution 4.0 International License.

METHODS

Ethical Considerations

This study was approved by the Bursa City Hospital Scientific Researches Ethics Committee (Date: 25.06.2025, Decision No: 2025-13/8). All data were anonymized, and the study was conducted in accordance with the Declaration of Helsinki.

A detailed flowchart of the study cohort selection is presented in **Figure 1**.



Note: None of the included patients received low-dose aspirin prophylaxis during pregnancy.
Abbreviations: PE, preeclampsia; Ob, obesity.

Figure 1. Study flowchart. Of 4,500 deliveries screened, 1,900 were excluded [missing first-trimester BMI (n=1,500), delivery at other institutions (n=300), major fetal anomalies (n=50), multiple pregnancies (n=50)], leaving 2,600 eligible records. From these, 400 women were selected by computer-generated simple random sampling to form four balanced groups (100 per group) without matching: PE(-) Ob(-), PE(+) Ob(-), PE(-) Ob(+), and PE(+) Ob(+). The analytic cohort comprises complete cases for variables used in the models. None of the included patients received low-dose aspirin prophylaxis.

BMI: Body-mass index, PE: Preeclampsia, Ob: Obesity

Study Design and Population

This retrospective, observational, and descriptive cohort study was conducted at the Department of Obstetrics and Gynecology, Bursa City Hospital. Medical records of pregnant women who gave birth between January 1, 2021, and May 1, 2023, were reviewed through the hospital's electronic medical record system. No medical interventions were performed as part of this study.

Pregnant women were included if they:

- Delivered at Bursa City Hospital within the specified study period,
- Had a singleton pregnancy resulting in a live birth,
- Delivered at or beyond 24 weeks of gestation,
- Had documented maternal height and weight in the first trimester (≤ 14 weeks),
- Had recorded maternal weight at the time of delivery,
- Had clear documentation of preeclampsia diagnosis (present or absent) with complete obstetric and neonatal outcome data, defined according to the American College of Obstetricians and Gynecologists (ACOG) criteria.¹⁰

Exclusion criteria were:

- Multiple pregnancies,
- Major fetal anomalies,
- Pre-existing systemic diseases (e.g., chronic kidney disease, connective tissue diseases, rheumatological conditions, malignancies),
- Missing or incomplete data.

A total of 4,500 deliveries were initially screened. After excluding women with missing first-trimester BMI data (n=1,500), deliveries at other institutions (n=300), major fetal anomalies (n=50), and multiple pregnancies (n=50), 2,600 women were eligible for analysis. From these eligible records, 100 women were randomly selected for each of the four predefined groups-non-obese without preeclampsia [PE(-) Ob(-)], non-obese with preeclampsia [PE(+) Ob(-)], obese without preeclampsia [PE(-) Ob(+)], and obese with preeclampsia [PE(+) Ob(+)]-to ensure balanced group sizes. Random sampling was implemented using computer-generated simple random selection; no covariate matching or weighting was used. This random selection process is acknowledged as a potential source of selection bias and is discussed as a study limitation.

None of the included patients received low-dose aspirin for preeclampsia prophylaxis during pregnancy, reflecting local clinical practice during the study period.

Grouping

Participants were classified into four groups according to the presence of preeclampsia (PE) and obesity (Ob):

- PE(-) Ob(-)
- PE(+) Ob(-)
- PE(-) Ob(+)
- PE(+) Ob(+)

Obesity was defined as a first-trimester BMI ≥ 30 kg/m², calculated using recorded height and weight. Gestational weight gain (GWG) was calculated as the difference between delivery weight and first-trimester weight and categorized according to the Institute of Medicine (IOM, 2009) guidelines as inadequate, appropriate, or excessive.¹¹

Data Collection

Data were extracted on:

- Demographic variables:
- Maternal age (years)
- Parity (number)
- BMI (kg/m²), first trimester
- GWG (kg)
- Chronic hypertension (yes)
- Gestational hypertension (yes)
- Pregestational diabetes (yes)
- Gestational diabetes (yes)

Obstetric variables:

- Gestational age at delivery (weeks)
- Mode of delivery (vaginal or cesarean)
- Preeclampsia diagnosis (yes/no)

Neonatal outcomes:

- Birth weight (grams)
- Low birth weight (LBW), defined as birth weight <2.500 g
- Macrosomia, defined as birth weight >4.000 g
- Preterm birth rates (<37 weeks and <34 weeks)
- 1st and 5th minute Apgar scores (numerical and categorized as <7)
- NICU admission and its indications (multiple indications per infant were allowed)

Restricted cubic spline (RCS) modeling: GWG was modeled as a continuous exposure using additive RCS terms. We used five knots with boundary knots at the 5th and 95th percentiles and interior knots at the 35th and 65th percentiles of the cohort GWG distribution (3.09, 9.36, 11.15, 12.90, and 17.20 kg). Binary outcomes (preterm <37 weeks; Apgar <7 at 1 and 5 minutes) were fit with generalized linear models (binomial family, logit link), and birth weight with linear regression. Each model included the four-level PE-obesity group as a main effect (no GWG×group interaction) and was adjusted for maternal age, smoking status, and IVF conception. Splines were centered at the cohort median GWG (11.15 kg), which served as the reference for plotting; predictions were displayed over the observed 5th-95th GWG range to limit extrapolation. Pointwise 95% confidence intervals were derived from model-based standard errors. We evaluated nonlinearity of GWG (RCS vs linear) and, in sensitivity analyses, the GWG×group interaction using likelihood-ratio tests; statistical significance was defined as two-sided $p < 0.05$.

Sample size and power: NICU admission was the prespecified primary outcome. Based on prior reports,²⁴ we anticipated NICU rates of ~40-45% in the PE(+) Ob(+) group versus ~20-25% in the PE(-) Ob(-) group. Using a four-group Pearson χ^2 comparison with a medium effect size (Cohen's $w = 0.25$), two-sided $\alpha = 0.05$, and 80% power, the required total sample size was ≈ 180 (≈ 45 per group). Our realized sample ($n = 400$; 100 per group) exceeded this a-priori target, providing adequate power for the primary comparisons and improved precision for adjusted analyses.

Covariate selection and model diagnostics: Covariates were selected a priori as a minimal sufficient adjustment set based on clinical plausibility and prior literature: maternal age, smoking, IVF conception, and gestational weight gain (GWG) category. Because parity may confound associations between exposure groups and perinatal outcomes, we conducted a sensitivity analysis additionally adjusting for parity. In a second sensitivity analysis, we refit the logistic models for LBW and for Apgar <7 at 1 and 5 minutes, as well as the model for NICU admission, additionally adjusting for gestational age at delivery (weeks); all other covariates were kept unchanged. Multicollinearity was assessed via variance inflation factors (VIF). Model calibration was evaluated using the Hosmer-Lemeshow goodness-of-fit test (10 risk groups), and discrimination using the area under the receiver-operating-characteristic curve (AUC). For each outcome we report N, events, Hosmer-Lemeshow p , AUC, and Nagelkerke's R^2 (Table 1). Adjusted odds ratios (aORs) from the main model and from sensitivity models additionally adjusting for parity, diabetes, and chronic hypertension; and for gestational age at delivery are provided in Table 2, and 3.

Statistical Analysis

All data analyses were performed using SPSS version 26 (IBM Corp., Armonk, NY, USA). Descriptive statistics are presented as means±standard deviation or medians (interquartile range) for continuous variables, and frequencies (%) for categorical variables. Comparisons among the four PE-obesity groups were conducted using the Chi-square test for categorical variables and ANOVA or Kruskal-Wallis test for continuous variables, depending on distribution; pairwise comparisons for non-normally distributed variables used appropriate post hoc tests with multiplicity control.

Associations between the PE-Obesity groups and perinatal outcomes (preterm birth, LBW, macrosomia, NICU admission, low Apgar scores) were analyzed using unadjusted and adjusted logistic regression models, with the PE(-) Ob(-) group as the reference category. Adjusted models controlled for potential confounders such as maternal age, smoking status, in vitro fertilization (IVF) conception, and GWG. Effect modification was evaluated by adding a multiplicative interaction term (PE×obesity) to each logistic regression model; PE and Ob were coded as binary indicators (1=yes, 0=no).

Table 1. Model diagnostics for adjusted logistic regression models

Outcome	n	Events	Hosmer-Lemeshow p	AUC	Nagelkerke R ²	VIF (min-max)
Preterm <37 weeks	400	82	0.166	0.74	0.20	1.10-2.78
Preterm <34 weeks	400	23	0.775	0.75	0.12	1.10-2.78
LBW (<2500 g)	400	71	0.585	0.81	0.30	1.10-2.78
NICU admission	400	50	0.767	0.73	0.14	1.10-2.78
Apgar <7 at 1 min	400	46	0.971	0.71	0.11	1.10-2.78
Apgar <7 at 5 min	400	25	0.561	0.70	0.08	1.10-2.78

Model diagnostics for each outcome in adjusted logistic regression models (base adjustment: age, smoking, IVF, GWG category). Hosmer-Lemeshow (10 groups), AUC, and Nagelkerke R² are reported. VIF range refers to the minimum and maximum variance inflation factors across adjustment covariates
AUC: Area under the curve, VIF: Variance inflation factor, LBW: Low birth weight, NICU: Neonatal intensive care unit, Apgar: Appearance, pulse, grimace, activity, respiration, IVF: In vitro fertilization, GWG: gestational weight gain

Table 2. Adjusted associations of PE-obesity group with perinatal outcomes

Outcome	Model	PE(+) Ob(-) vs ref	PE(-) Ob(+) vs ref	PE(+) Ob(+) vs ref
Preterm <37 weeks	Base (age, smoking, IVF, GWG)	7.64 (2.78-20.95)	2.03 (0.64-6.40)	12.17 (4.36-33.97)
Preterm <37 weeks	Sensitivity (+parity, diabetes, chronic hypertension)	7.68 (2.80-21.09)	2.02 (0.64-6.38)	12.14 (4.35-33.91)
Preterm <34 weeks	Base (age, smoking, IVF, GWG)	8.04 (0.96-67.59)	3.70 (0.37-37.05)	18.44 (2.24-151.86)
Preterm <34 weeks	Sensitivity (+parity, diabetes, chronic hypertension)	8.11 (0.96-68.35)	3.71 (0.37-37.13)	18.43 (2.24-151.85)
LBW (<2500 g)	Base (age, smoking, IVF, GWG)	11.23 (3.67-34.35)	0.61 (0.11-3.51)	19.61 (6.22-61.81)
LBW (<2500 g)	Sensitivity (+parity, diabetes, chronic hypertension)	11.04 (3.60-33.85)	0.62 (0.11-3.52)	19.99 (6.32-63.22)
NICU admission	Base (age, smoking, IVF, GWG)	5.46 (1.51-19.80)	2.76 (0.68-11.23)	12.41 (3.45-44.62)
NICU admission	Sensitivity (+parity, diabetes, chronic hypertension)	5.77 (1.58-21.09)	2.73 (0.67-11.16)	12.45 (3.44-45.08)
Apgar <7 at 1 min	Base (age, smoking, IVF, GWG)	3.54 (1.09-11.46)	2.17 (0.60-7.85)	8.54 (2.67-27.33)
Apgar <7 at 1 min	Sensitivity (+parity, diabetes, chronic hypertension)	3.66 (1.12-11.90)	2.16 (0.59-7.83)	8.52 (2.65-27.36)
Apgar <7 at 5 min	Base (age, smoking, IVF, GWG)	3.07 (0.59-15.87)	2.23 (0.39-12.83)	7.86 (1.62-38.15)
Apgar <7 at 5 min	Sensitivity (+parity, diabetes, chronic hypertension)	3.14 (0.60-16.35)	2.21 (0.38-12.75)	7.82 (1.61-38.05)

Adjusted odds ratios (aOR; 95% CI) comparing PE-Ob groups to the reference PE(-) Ob(-). Two models are shown: base (age, smoking, IVF, GWG) and sensitivity (+parity, diabetes, chronic hypertension). PE: Preeclampsia, Ob: Obesity, IVF: In vitro fertilization, GWG: Gestational weight gain, LBW: Low birth weight, NICU: Gestational weight gain, Apgar: Appearance, pulse, grimace, activity, respiration, aOR: Adjusted odds ratio, CI: Confidence interval

Table 3. Sensitivity analysis adjusting for gestational age at delivery

Outcome	Comparison	Events/n	aOR	95% CI	p-value
LBW (<2.500 g)	PE(+) Ob(-) vs PE(-) Ob(-)	28/100	8.31	2.65-26.04	0.000
LBW (<2.500 g)	PE(-) Ob(+) vs PE(-) Ob(-)	2/100	0.57	0.10-3.32	0.536
LBW (<2.500 g)	PE(+) Ob(+) vs PE(-) Ob(-)	37/100	12.06	3.65-39.90	0.000
Apgar <7 at 1 min	PE(+) Ob(-) vs PE(-) Ob(-)	13/100	1.40	0.38-5.11	0.610
Apgar <7 at 1 min	PE(-) Ob(+) vs PE(-) Ob(-)	7/100	1.54	0.38-6.28	0.550
Apgar <7 at 1 min	PE(+) Ob(+) vs PE(-) Ob(-)	22/100	2.21	0.58-8.50	0.247
Apgar <7 at 5 min	PE(+) Ob(-) vs PE(-) Ob(-)	6/100	1.36	0.24-7.63	0.725
Apgar <7 at 5 min	PE(-) Ob(+) vs PE(-) Ob(-)	4/100	1.72	0.28-10.66	0.560
Apgar <7 at 5 min	PE(+) Ob(+) vs PE(-) Ob(-)	13/100	2.26	0.39-13.13	0.363
NICU admission	PE(+) Ob(-) vs PE(-) Ob(-)	15/100	2.62	0.67-10.17	0.164
NICU admission	PE(-) Ob(+) vs PE(-) Ob(-)	7/100	2.13	0.49-9.32	0.313
NICU admission	PE(+) Ob(+) vs PE(-) Ob(-)	25/100	4.22	1.05-17.00	0.043

aOR for group comparisons versus PE(-) Ob(-), from logistic regression models that adjust for maternal age, smoking, IVF, GWG category, and gestational age at delivery (weeks). aOR: Adjusted odds ratio, CI: Confidence interval, LBW: Low birth weight, PE: Preeclampsia, Ob: Obesity, Apgar: Appearance, pulse, grimace, activity, respiration, NICU: Gestational weight gain

RESULTS

Baseline Characteristics

Baseline maternal and clinical characteristics according to the Preeclampsia-Obesity (PE-Ob) groups are summarized in [Table 4](#). Maternal age and BMI differed significantly across groups ($p<0.001$).

Women in the PE(+) Ob(+) group had the highest median age (32 years) compared to other groups. BMI was significantly higher in the obese groups, with median (Q1-Q3) values of 33.5 (31.78-37.32) in the PE(-) Ob(+) group and 34.25 (32.10-38.68) in the PE(+) Ob(+) group. Gestational weight gain (GWG) showed a decreasing trend from the PE(-) Ob(-) to PE(+) Ob(+) group (median: 13.25 vs. 8.35 kg, $p<0.001$).

There were no significant differences in gravida, abortions, smoking status, IVF conception, or diabetes status among groups, except for GWG category distributions ($p<0.001$) which showed higher proportions of low GWG in preeclamptic

groups. In the PE(+) groups, the proportions of early-onset PE (<34 weeks) and late-onset PE (≥ 34 weeks) were 23% vs 77% in the PE(+) Ob(-) group, and 37% vs 63% in the PE(+) Ob(+) group, respectively ($p=0.044$).

Perinatal Outcomes

Perinatal outcomes according to PE-Ob groups are presented in [Table 5](#).

The PE(+) Ob(+) group had the lowest median gestational age at delivery [37.0 (35.0-38.0) weeks], significantly lower than PE(-) Ob(-) ($p<0.001$).

Mean birth weight was also significantly lower in the PE(+) Ob(+) group (2610 $\pm$ 573 g) compared to the PE(-) Ob(-) group (3185 $\pm$ 453 g, $p<0.001$).

Table 4. Baseline demographic and clinical characteristics according to preeclampsia-obesity groups

Variable	PE(-) Ob(-) (n=100)	PE(+) Ob(-) (n=100)	PE(-) Ob(+) (n=100)	PE(+) Ob(+) (n=100)	p-value
Maternal age	28 (26-30) ^a	28.5 (26-31.25) ^a	29.5 (27-33) ^b	32 (29-36) ^c	<0.001 ¹
BMI	24.0 (22.58-25.53) ^a	24.1 (23.18-25.13) ^a	33.5 (31.78-37.32) ^b	34.25 (32.10-38.68) ^b	<0.001 ¹
Gravida	2.7±1.1	2.6±1.1	2.7±1.1	2.7±1.1	0.720 ²
Abortions	0.6±0.7	0.7±0.8	0.6±0.7	0.6±0.8	0.860 ²
Smoking, n (%)	8 (8.0%)	8 (8.0%)	15 (15.0%)	10 (10.0%)	0.310 ³
IVF, n (%)	9 (9.0%)	8 (8.0%)	10 (10.0%)	9 (9.0%)	0.970 ³
Chronic hypertension (yes), n (%)	0 (0.0%)	2 (2.0%)	3 (3.00%)	5 (5.0%)	0.149 ³
Pregestational diabetes (yes), n (%)	0 (0.0%)	1 (1.0%)	1 (1.00%)	2 (2.0%)	0.568 ³
Gestational diabetes (yes), n (%)	1 (1.0%)	2 (2.0%)	3 (3.00%)	4 (4.0%)	0.562 ³
Early-onset PE (<34 weeks), n (%)	N/A	23 (23%)	N/A	37 (37%)	0.044 ³
Late-onset PE (≥34 weeks), n (%)	N/A	77 (77%)	N/A	63 (63%)	0.044 ³
GWG	13.25 (11.38-15.22) ^c	11.35 (8.50-13.62) ^b	10.10 (6.78-13.30) ^b	8.35 (5.90-11.95) ^a	<0.001 ³
Low GWG, n (%)	14 (14.0%) ^a	38 (38.0%) ^b	17 (17.0%) ^a	24 (24.0%) ^a	<0.001 ³
Normal GWG, n (%)	51 (51.0%) ^b	36 (36.0%) ^a	26 (26.0%) ^a	32 (32.0%) ^a	0.002 ³
High GWG, n (%)	35 (35.0%) ^a	26 (26.0%) ^a	57 (57.0%)	44 (44.0%)	<0.001 ³

Data are presented as median (IQR), mean±standard deviation or number (percentage). Different superscript letters indicate statistically significant pairwise comparisons within rows (p<0.05). Superscript letters are ranked from lowest to highest group values (c>b>a). Statistical comparisons were performed using the Kruskal-Wallis test, ANOVA, or Chi-square test, as appropriate. PE: Preeclampsia, Ob: Obesity, BMI: Body-mass index, GWG: Gestational weight gain. GWG categories were defined according to the 2009 Institute of Medicine (IOM) guidelines for singleton pregnancies: recommended total GWG ranges are 12.5-18 kg (BMI <18.5 kg/m²), 11.5-16 kg (BMI 18.5-24.9 kg/m²), 7-11.5 kg (BMI 25-29.9 kg/m²), and 5-9 kg (BMI ≥30 kg/m²).

Table 5. Comparison of perinatal outcomes according to preeclampsia-obesity groups

Variable	PE(-) Ob(-) (n=100)	PE(+) Ob(-) (n=100)	PE(-) Ob(+) (n=100)	PE(+) Ob(+) (n=100)	p-value
Gestational age at delivery (weeks)	39.0 (38.0-40.0) ^b	37.0 (36.0-38.0) ^a	38.5 (37.0-40.0) ^b	37.0 (35.0-38.0) ^a	<0.001 ¹
Birth weight (g)	3185±453 ^b	2716±541 ^a	3422±485 ^c	2610±573 ^a	<0.001 ²
Cesarean delivery, n (%)	32 (32.0%) ^a	57 (57.0%) ^b	49 (49.0%) ^b	77 (77.0%) ^c	<0.001 ³
Previous cesarean	17 (53.1%)	25 (43.9%)	22 (44.9%)	32 (41.6%)	
Arrest of labor/failure to progress	5 (15.6%)	8 (14.0%)	8 (16.3%)	10 (13.0%)	
Cephalopelvic disproportion	3 (9.4%)	10 (17.5%)	6 (12.2%)	12 (15.6%)	
Macrosomia	3 (9.4%)	0 (0%)	4 (8.2%)	0 (0%)	
Fetal distress	2 (6.3%)	4 (7.0%)	3 (6.1%)	8 (10.4%)	
Malpresentation	2 (6.3%)	4 (7.0%)	3 (6.1%)	8 (10.4%)	
Placenta previa	0 (0%)	2 (3.5%)	1 (2.0%)	3 (3.9%)	
Failed induction	0 (0%)	4 (7.0%)	2 (4.1%)	4 (5.2%)	
Vaginal delivery, n (%)	68 (68.0%) ^c	43 (43.0%) ^b	51 (51.0%) ^b	23 (23.0%) ^a	<0.001 ³
NICU admission, n (%)	3 (3.0%) ^a	15 (15.0%) ^b	7 (7.0%) ^{a,b}	25 (25.0%) ^b	<0.001 ³
Sepsis, n (%)		1 (7%)	1 (14%)		
RDS, n (%)		2 (13%)	1 (14%)	4 (16%)	
Jaundice, n (%)	1 (33%)	3 (20%)	1 (14%)	7 (28%)	
TTN, n (%)	1 (33%)	7 (47%)	2 (29%)	8 (32%)	
Hypoglycemia, n (%)	1 (33%)	2 (13%)	2 (29%)	5 (20%)	
IVH, n (%)				1 (4%)	
Preterm birth <37 weeks, n (%)	5 (5.0%) ^a	30 (30.0%) ^b	9 (9.0%) ^a	38 (38.0%) ^b	<0.001 ³
Preterm birth <34 weeks, n (%)	1 (1.0%) ^a	7 (7.0%) ^b	3 (3.0%) ^{a,b}	12 (12.0%) ^b	0.005 ³
Low birth weight <2500g, n (%)	4 (4.0%) ^a	28 (28.0%) ^b	2 (2.0%) ^a	37 (37.0%) ^b	<0.001 ³
Macrosomia >4000g, n (%)	5 (5.0%) ^a	1 (1.0%) ^a	9 (9.0%) ^b	1 (1.0%) ^a	0.009 ³
Apgar score <7 at 1 min, n (%)	4 (4.0%) ^a	13 (13.0%) ^b	7 (7.0%) ^{a,b}	22 (22.0%) ^b	<0.001 ³
Apgar score <7 at 5 min, n (%)	2 (2.0%) ^a	6 (6.0%) ^a	4 (4.0%) ^a	13 (13.0%) ^b	0.008 ³

Data are presented as median (interquartile range), mean±standard deviation, or number (percentage), as appropriate. Superscript letters indicate statistically significant pairwise comparisons within rows (p<0.05). Superscript letters are ranked from lowest to highest group values (c>b>a). Statistical comparisons were performed using the Kruskal-Wallis test, ANOVA, or Chi-square test, as appropriate. PE: Preeclampsia, Ob: Obesity, NICU: Neonatal intensive care unit, RDS: Respiratory distress syndrome, TTN: Transient tachypnea of the newborn, IVH: intraventricular hemorrhage, Apgar: Appearance, pulse, grimace, activity, respiration. "Note: Infants could have >1 indication; counts by indication may exceed the number of NICU admissions."

The cesarean delivery rate was highest in the PE(+) Ob(+) group (77%) ($p<0.001$). The primary indications for cesarean delivery were previous cesarean section (53.1%, 43.9%, 44.9%, and 41.6% in the PE(-) Ob(-), PE(+) Ob(-), PE(-) Ob(+), and PE(+) Ob(+) groups, respectively), arrest of labor or failure to progress (15.6%, 14.0%, 16.3%, and 13.0%), cephalopelvic disproportion (9.4%, 17.5%, 12.2%, and 15.6%), fetal distress (6.3%, 7.0%, 6.1%, and 10.4%), malpresentation (6.3%, 7.0%, 6.1%, and 10.4%), placenta previa (0%, 3.5%, 2.0%, and 3.9%), failed induction (0%, 7.0%, 4.1%, and 5.2%), and macrosomia (9.4%, 0%, 8.2%, and 0%, respectively).

NICU admission was more common in the PE(+) Ob(+) (25%) and PE(+) Ob(-) (15%) groups than in the reference group (3%) ($p<0.001$).

Rates of preterm birth <37 weeks, preterm birth <34 weeks, and low birth weight (<2500 g) were significantly higher in the PE(+) groups ($p<0.001$).

Macrosomia (>4000 g) was more prevalent in the PE(-) Ob(+) group (9%) compared to other groups ($p=0.009$).

Apgar scores <7 at 1 and 5 minutes were also more frequent in the PE(+) Ob(+) group ($p<0.001$).

NICU indications included respiratory distress syndrome (RDS), transient tachypnea of the newborn (TTN), jaundice, hypoglycemia, sepsis, and intraventricular hemorrhage (IVH). Infants could have multiple indications; TTN, RDS, and jaundice were most common, with higher frequencies in PE(+) Ob(+).

Univariable Unadjusted Risk Estimates

Univariable logistic regression results are shown in [Table 6](#). Compared to the PE(-) Ob(-) group, the combined PE(+) Ob(+) group had significantly increased odds for preterm birth <37 weeks [OR: 11.65 (95% CI: 4.35-31.21)], preterm birth <34 weeks [OR: 13.5 (1.72-105.94)], LBW [OR: 14.1 (4.79-41.48)], NICU admission [OR: 10.78 (3.13-37.06)], and Apgar score <7 at 1 min [OR: 6.77 (2.24-20.47)], and Apgar score <7 at 5 min [OR: 7.32 (1.61-33.36)] (all $p<0.05$).

The PE(+) Ob(-) group also showed elevated risks for these outcomes except for preterm birth <34 weeks and Apgar score <7 at 5 min, while the PE(-) Ob(+) group, representing obesity alone, did not significantly increase the risk for preterm birth, LBW, or Apgar score <7 at 1 min.

Multivariable Adjusted Risk Estimates

After adjusting for maternal age, smoking status, IVF conception, and GWG category, the multivariable logistic regression analysis showed similar associations ([Table 7](#)). The PE(+) Ob(+) group had the highest risk for preterm birth <37 weeks [adjusted OR: 12.17 (95% CI: 4.36-33.97)], preterm birth <34 weeks [adjusted OR: 18.44 (2.24-151.86)], low birth weight [adjusted OR: 19.61 (6.22-61.81)], NICU admission [adjusted OR: 12.41 (3.45-44.62)], Apgar score <7 at 1 min [adjusted OR: 8.54 (2.67-27.33)], and Apgar score <7 at 5 min [adjusted OR: 7.86 (1.62-38.15)] (all $p<0.05$).

The PE(+) Ob(-) group also showed elevated risks for these adverse outcomes, except for preterm birth <34 weeks

Table 6. Univariable logistic regression analysis of preeclampsia-obesity groups for adverse perinatal outcomes

Outcome	Group vs Ref PE(-) Ob(-)	Events, n/N vs ref	OR (95% CI)	p-value
Preterm birth <37 weeks	PE(+) Ob(+)	38/100 vs 5/100	11.65 (4.35-31.21)	<0.001
Preterm birth <37 weeks	PE(+) Ob(-)	30/100 vs 5/100	8.14 (3.01-22.04)	<0.001
Preterm birth <37 weeks	PE(-) Ob(+)	9/100 vs 5/100	1.88 (0.61-5.82)	0.274
Preterm birth <34 weeks	PE(+) Ob(+)	12/100 vs 1/100	13.5 (1.72-105.94)	0.013
Preterm birth <34 weeks	PE(+) Ob(-)	7/100 vs 1/100	7.45 (0.9-61.73)	0.063
Preterm birth <34 weeks	PE(-) Ob(+)	3/100 vs 1/100	3.06 (0.31-29.95)	0.336
Low birth weight (<2500 g)	PE(+) Ob(+)	37/100 vs 4/100	14.1 (4.79-41.48)	<0.001
Low birth weight (<2500 g)	PE(+) Ob(-)	28/100 vs 4/100	9.33 (3.13-27.8)	<0.001
Low birth weight (<2500 g)	PE(-) Ob(+)	2/100 vs 4/100	0.49 (0.09-2.74)	0.416
Macrosomia (>4000 g)	PE(+) Ob(+)	1/100 vs 5/100	0.19 (0.02-1.67)	0.135
Macrosomia (>4000 g)	PE(+) Ob(-)	1/100 vs 5/100	0.19 (0.02-1.67)	0.135
Macrosomia (>4000 g)	PE(-) Ob(+)	9/100 vs 5/100	1.88 (0.61-5.82)	0.274
NICU admission	PE(+) Ob(+)	25/100 vs 3/100	10.78 (3.13-37.06)	<0.001
NICU admission	PE(+) Ob(-)	15/100 vs 3/100	5.71 (1.6-20.39)	0.007
NICU admission	PE(-) Ob(+)	7/100 vs 3/100	2.43 (0.61-9.69)	0.207
Apgar score <7 at 1 min	PE(+) Ob(+)	22/100 vs 4/100	6.77 (2.24-20.47)	<0.001
Apgar score <7 at 1 min	PE(+) Ob(-)	13/100 vs 4/100	3.59 (1.13-11.41)	0.031
Apgar score <7 at 1 min	PE(-) Ob(+)	7/100 vs 4/100	1.81 (0.51-6.38)	0.358
Apgar score <7 at 5 min	PE(+) Ob(+)	13/100 vs 2/100	7.32 (1.61-33.36)	0.010
Apgar score <7 at 5 min	PE(+) Ob(-)	6/100 vs 2/100	3.13 (0.62-15.89)	0.169
Apgar score <7 at 5 min	PE(-) Ob(+)	4/100 vs 2/100	2.04 (0.37-11.41)	0.416

Reference group is PE(-) Ob(-). PE: Preeclampsia, Ob: Obesity, OR: Odds ratio, CI: Confidence interval, NICU: Neonatal intensive care unit, Apgar: Appearance, pulse, grimace, activity, respiration

Table 7. Adjusted odds ratios for preeclampsia-obesity groups (adjusted for maternal age, smoking, IVF and GWG category)

Outcome	Group vs Ref PE(-) Ob(-)	Events, n/N vs ref	Adjusted OR (95% CI)	p-value
Preterm birth <37 weeks	PE(+) Ob(+)	38/100 vs 5/100	12.17 (4.36-33.97)	<0.001
Preterm birth <37 weeks	PE(+) Ob(-)	30/100 vs 5/100	7.64 (2.78-20.95)	<0.001
Preterm birth <37 weeks	PE(-) Ob(+)	9/100 vs 5/100	2.03 (0.64-6.4)	0.227
Preterm birth <34 weeks	PE(+) Ob(+)	12/100 vs 1/100	18.44 (2.24-151.86)	0.007
Preterm birth <34 weeks	PE(+) Ob(-)	7/100 vs 1/100	8.04 (0.96-67.59)	0.055
Preterm birth <34 weeks	PE(-) Ob(+)	3/100 vs 1/100	3.7 (0.37-37.05)	0.265
Low birth weight (<2500 g)	PE(+) Ob(+)	37/100 vs 4/100	19.61 (6.22-61.81)	<0.001
Low birth weight (<2500 g)	PE(+) Ob(-)	28/100 vs 4/100	11.23 (3.67-34.35)	<0.001
Low birth weight (<2500 g)	PE(-) Ob(+)	2/100 vs 4/100	0.61 (0.11-3.51)	0.583
Macrosomia (>4000 g)	PE(+) Ob(+)	1/100 vs 5/100	0.14 (0.02-1.36)	0.090
Macrosomia (>4000 g)	PE(+) Ob(-)	1/100 vs 5/100	0.19 (0.02-1.73)	0.141
Macrosomia (>4000 g)	PE(-) Ob(+)	9/100 vs 5/100	1.49 (0.45-4.96)	0.515
NICU admission	PE(+) Ob(+)	25/100 vs 3/100	12.41 (3.45-44.62)	<0.001
NICU admission	PE(+) Ob(-)	15/100 vs 3/100	5.46 (1.51-19.8)	0.010
NICU admission	PE(-) Ob(+)	7/100 vs 3/100	2.76 (0.68-11.23)	0.155
Apgar score <7 at 1 min	PE(+) Ob(+)	22/100 vs 4/100	8.54 (2.67-27.33)	<0.001
Apgar score <7 at 1 min	PE(+) Ob(-)	13/100 vs 4/100	3.54 (1.09-11.46)	0.035
Apgar score <7 at 1 min	PE(-) Ob(+)	7/100 vs 4/100	2.17 (0.6-7.85)	0.238
Apgar score <7 at 5 min	PE(+) Ob(+)	13/100 vs 2/100	7.86 (1.62-38.15)	0.011
Apgar score <7 at 5 min	PE(+) Ob(-)	6/100 vs 2/100	3.07 (0.59-15.87)	0.182
Apgar score <7 at 5 min	PE(-) Ob(+)	4/100 vs 2/100	2.23 (0.39-12.83)	0.369

Multivariable logistic regression models were adjusted for maternal age, smoking status, in vitro fertilization IVF and GWG category. Reference group is PE(-) Ob(-). IVF: In vitro fertilization, GWG: gestational weight gain, PE: Preeclampsia, Ob: Obesity, OR: Odds ratio, CI: Confidence interval, NICU: Neonatal intensive care unit, Apgar: Appearance, pulse, grimace, activity, respiration

and Apgar score <7 at 5 min, which were not statistically significant. In contrast, the PE(-) Ob(+) group, representing obesity alone, was not significantly associated with increased risk for any of these perinatal outcomes in the multivariable adjusted analysis.

Interaction Analysis

For preterm birth <37 weeks, the PE×obesity interaction was $\beta = -0.24$ (SE 0.66), corresponding to OR_{int} = 0.79 (95% CI 0.22-2.85), $p = 0.713$, indicating no significant multiplicative interaction. Similar null interactions were observed for other outcomes: preterm <34 weeks (OR_{int} 0.62, 95% CI 0.05-7.52, $p = 0.707$), low birth weight (OR_{int} 2.84, 95% CI 0.45-17.84, $p = 0.265$), NICU admission (OR_{int} 0.82, 95% CI 0.17-3.95, $p = 0.807$), Apgar <7 at 1 min (OR_{int} 1.11, 95% CI 0.25-4.92, $p = 0.888$), and Apgar <7 at 5 min (OR_{int} 1.15, 95% CI 0.15-8.59, $p = 0.891$). These findings suggest that the markedly higher risks in the PE(+) Ob(+) group are primarily explained by the main effects rather than a multiplicative interaction.

Model Diagnostics and Sensitivity Analyses

Across adjusted logistic models, calibration was acceptable (Hosmer-Lemeshow $p = 0.166$ -0.971) and discrimination was fair (AUC 0.70-0.81); Nagelkerke's R^2 ranged from 0.08 to 0.30. Multicollinearity was not concerning (VIF across adjustment covariates 1.10-2.78) (Table 5). Insensitivity models additionally adjusting for parity, diabetes, and chronic hypertension, group-specific aORs were essentially unchanged; for example,

for preterm birth <37 weeks the aOR for PE(+) Ob(+) versus PE(-) Ob(-) was 12.17 (95% CI 4.36-33.97) in the base model and 12.14 (4.35-33.91) with the additional adjustment, and for low birth weight the corresponding aORs were 19.61 (6.22-61.81) and 19.99 (6.32-63.22), respectively (Table 6). Results from the models additionally adjusted for parity, diabetes, and chronic hypertension were unchanged. Results of the parity-adjusted models were unchanged. With additional adjustment for gestational age at delivery, associations with LBW in the PE(+) groups remained robust (PE(+) Ob(-) aOR 8.31; 95% CI 2.65-26.04 and PE(+) Ob(+) aOR 12.06; 95% CI 3.65-39.90), whereas effects for Apgar <7 at 1 and 5 minutes were attenuated and lost statistical significance; for NICU admission, only the PE(+) Ob(+) group remained significant (aOR 4.22; 95% CI 1.05-17.00) (Table 7).

Forest Plot and Spline Analyses

The forest plot in Figure 2 visualizes the univariable (model 1) and multivariable adjusted (model 2) odds ratios for significant perinatal outcomes across PE-Ob groups. The combined PE(+) Ob(+) group showed the highest adjusted risks for preterm birth (<37 and <34 weeks), low birth weight, NICU admission, and low Apgar scores at both 1 and 5 minutes.

Figure 3 displays adjusted restricted cubic-spline (RCS) curves for gestational weight gain (GWG) modeled as a continuous exposure. Additive models (no GWG×group interaction) included maternal age, smoking, and IVF, with the four-level PE-obesity group entered as a main effect. A five-knot

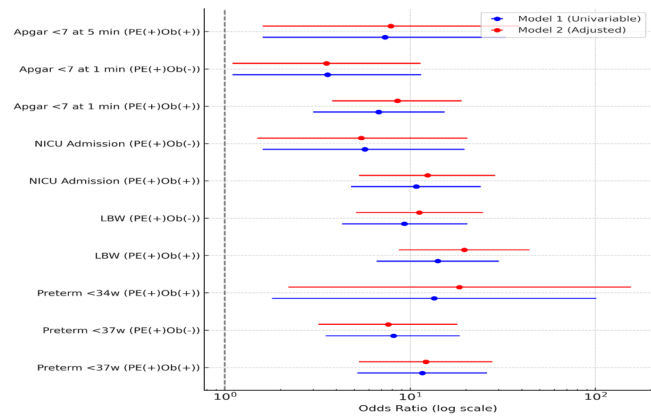


Figure 2. Forest plot comparing the (ORs) and 95% confidence intervals for significant perinatal outcomes across PE-Ob groups. Blue dots represent univariable (model 1) and red dots represent multivariable adjusted (model 2) logistic regression models. The PE(+) Ob(+) group showed the highest risks for preterm birth (<34 and <37 weeks), low birth weight, NICU admission, and Apgar scores <7 at 1 and 5 minutes.

ORs: Odds ratios, PE: Preeclampsia, Ob: Obesity, NICU: Neonatal intensive care unit, Apgar: Appearance, pulse, grimace, activity, respiration

RCS was used with knots at the 5th, 35th, 50th, 65th, and 95th percentiles of the cohort GWG distribution (3.09, 9.36, 11.15, 12.90, and 17.20 kg); curves are shown over the observed 5th-95th GWG range. The cohort median (11.15 kg) is indicated by a dashed grey vertical reference line, and small triangle markers denote knot locations; shaded bands represent pointwise 95% CIs. In the PE(+) Ob(+) group, predicted risk of preterm birth <37 weeks increased as GWG decreased below roughly 10-12 kg and rose modestly again at higher GWG, with broadly similar-though flatter-patterns in other groups and wider CIs toward the tails. For Apgar <7 at 1 min, there was modest curvature (nonlinearity p=0.029), whereas for Apgar <7 at 5 min no consistent association across the GWG range was evident (p=0.274). For birth weight, higher GWG was associated with higher adjusted mean birth weight across the range (nonlinearity p=0.333), and predictions for the PE(+) Ob(+) group remained lower across GWG with greater uncertainty near the extremes. These descriptive patterns align with formal tests that did not support retaining a GWG×group interaction, so additive models are presented.

Additionally, **Table 8** presents detailed perinatal outcomes stratified by BMI subgroups within the PE(+) Ob(+) group. This exploratory subgroup analysis suggested a possible trend of increased risk for preterm birth <37 weeks, NICU

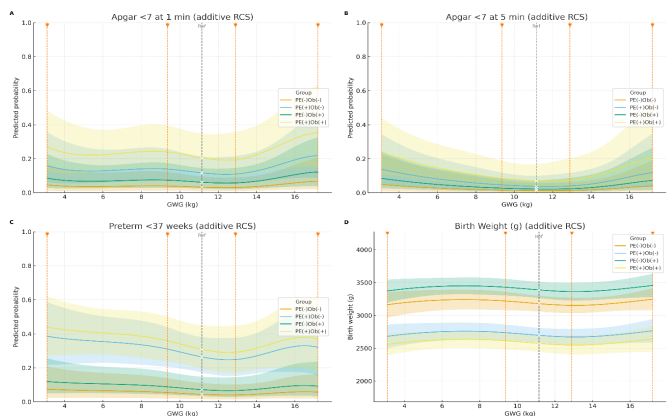


Figure 3. Adjusted RCS curves for GWG across PE-Ob groups. Additive models (no GWG×group interaction) included maternal age, smoking, and IVF, with the four-level PE-Ob group entered as a main effect. A five-knot RCS was used with boundary knots at the cohort 5th and 95th percentiles and interior knots at the 35th and 65th percentiles (3.09, 9.36, 11.15, 12.90, and 17.20 kg; the 11.15 kg cohort median is shown as a dashed grey reference line). Curves are displayed over the observed 5th-95th GWG range; shaded bands are pointwise 95% confidence intervals; small triangles mark knot locations.

RCS: Restricted cubic-spline, GWG: Gestational weight gain, PE: Preeclampsia, Ob: Obesity, IVF: In vitro fertilization

admission, and low Apgar scores with higher BMI, while no significant trend was observed for LBW, macrosomia, preterm birth <34 weeks, or median GWG.

DISCUSSION

In this study, the co-occurrence of preeclampsia and obesity was associated with more unfavorable perinatal outcomes than either condition alone. The PE(+) Ob(+) group had the highest odds of preterm birth (<37 and <34 weeks), low birth weight, NICU admission, low Apgar scores, and cesarean delivery, and also showed the highest rate of early-onset preeclampsia. Analysis of cesarean indications showed that prior cesarean section was the most frequent reason across all groups, followed by arrest of labor/failure to progress, cephalopelvic disproportion, fetal distress, malpresentation, placenta previa, failed induction, and macrosomia; accordingly, the higher overall cesarean rate in PE(+) Ob(+) appears largely attributable to prior uterine surgery and labor complications rather than obesity per se. After adjustment for maternal age, smoking status, IVF, and GWG, these associations remained significant. Importantly, the prespecified PE×obesity interaction was not statistically significant, suggesting that the excess risk with co-occurrence

Table 8. Perinatal outcomes by BMI subgroups within the PE(+) Ob(+) group					
Variable	BMI (30-<35) n=53	BMI (35-<40) n=22	BMI (40-<50) n=18	BMI (≥50) n=7	p for trend
Median GWG (kg)	9.0	7.45	8.35	11.5	0.60
Preterm birth <37 weeks (%)	18(34.0%)	8 (36.4%)	9 (50.0%)	4 (57.1%)	<0.001
Preterm birth <34 weeks (%)	7 (13.2%)	2 (9.1%)	2 (11.1%)	2 (28.6%)	0.60
LBW (%)	22 (41.5%)	5 (22.7%)	6 (33.3%)	4 (57.1%)	0.60
Macrosomia (%)	7 (13.2%)	2 (9.1%)	3 (16.6%)	2 (28.6%)	0.20
NICU admission (%)	8 (15.1%)	5 (22.7%)	9 (50.0%)	4 (57.1%)	<0.001
Apgar score <7 at 1 min (%)	7 (13.2%)	4 (18.2%)	8 (44.4%)	4 (57.1%)	<0.001
Apgar score <7 at 5 min (%)	4 (7.5%)	4 (18.2%)	5 (27.8%)	0.0	0.80

Data are presented as counts (n), median (GWG) or percentages (%). Trend p-values were calculated using Spearman rank correlation across increasing BMI subgroups. Results for BMI ≥50 should be interpreted with caution due to small sample size.
BMI: Body-mass index, PE: Preeclampsia, Ob: Obesity, GWG: Gestational weight gain, LBW: Low birth weight, NICU: Neonatal intensive care unit

is chiefly explained by main effects-particularly preeclampsia-rather than multiplicative effect modification. In contrast, obesity alone [PE(-) Ob(+)] was not significantly associated with these adverse outcomes. In additive restricted cubic spline analyses of GWG, curves showed a broadly U-shaped pattern with the lowest predicted risk around mid-range GWG (~10-12 kg) and widening uncertainty at the extremes. Across groups, lower GWG was associated with higher risk of preterm birth and with lower predicted birth weight; absolute risks were consistently highest in PE(+) Ob(+), reflecting vertical shifts rather than qualitatively different curve shapes. Evidence of nonlinearity was modest and outcome-specific (e.g., present for Apgar <7 at 1 minute, not for Apgar <7 at 5 minutes or preterm <37 weeks), supporting presentation of additive rather than interaction models.

As stated in the literature, the presence of preeclampsia leads to preterm birth due to fetal-placental complications such as intrauterine growth restriction and abnormal Doppler flows, which in turn increases the need for NICU admission.^{5,12} In our study, the notably higher rates of preterm birth in the PE(+) groups support this mechanism. Especially the fact that the PE(+) Ob(+) group had the highest preterm birth rate suggests that obesity alone is not a significant risk factor for preterm birth but may exert a synergistic effect when combined with preeclampsia. This finding partly contrasts with some studies reporting that obesity is an independent risk factor for preterm birth during pregnancy.¹³⁻¹⁵ This highlights the importance of considering the combined effects rather than isolated risk factors when managing high-risk pregnancies.

Some previous studies have reported unexpectedly lower rates of low birth weight in moderately obese groups.¹⁶ In our study, the higher rate of macrosomia in the PE(-) Ob(+) group compared to the other groups appears to support this finding. However, logistic regression analyses did not show obesity as an independent significant risk factor for macrosomia. Therefore, this trend needs to be confirmed by future studies with larger sample sizes. On the other hand, some studies have defined obesity as a risk factor for low birth weight as well.^{17,18} Our results, on the other hand, show that obesity by itself did not raise the risk of having a low-birth-weight baby. This risk became much higher when paired with preeclampsia. The fact that both the PE(+) Ob(+) and PE(+) Ob(-) groups had a much higher chance of having a baby with low birth weight implies that obesity may amplify this effect when paired with preeclampsia.

Our study found that obesity alone did not significantly raise the risk of NICU admission, but obesity combined with preeclampsia did considerably raise the risk. This suggests that obesity by itself may not be an independent risk factor for NICU admission, but its combination with preeclampsia may substantially increase this risk. This contrasts with some previous studies that defined obesity as an independent risk factor for NICU admission^{5,19,20} and highlights the determining role of preeclampsia.

In our study, when Apgar scores were examined, the presence of preeclampsia increased the risk of low Apgar scores, and this risk was further elevated when preeclampsia coexisted with obesity. In contrast, obesity alone did not have a

significant effect on Apgar scores. This finding is consistent with literature reporting that preeclampsia can lower Apgar scores, and it indicates that preeclampsia may strengthen this negative effect when combined with obesity.²¹⁻²³ However, one study reported that in general patient groups with both preeclampsia and obesity, there was no significant association with low Apgar scores, whereas only in patients with BMI ≥40 and preeclampsia was the risk of low Apgar score found to be increased.²⁴ This suggests that in our study, even at lower levels of obesity, preeclampsia may play a determining role in amplifying its negative impact on Apgar scores.

In addition, the exploratory subgroup analysis within the PE(+) Ob(+) cohort demonstrated that increasing BMI levels were associated with higher rates of preterm birth <37 weeks, NICU admission, and low 1-minute Apgar scores. This trend was particularly evident in patients with BMI ≥40. However, no clear trend was observed for preterm birth <34 weeks, low birth weight, or macrosomia. These results, although limited by small sample sizes in the highest BMI strata, suggest that extreme obesity within preeclamptic women may further aggravate perinatal risks, especially for preterm birth and early neonatal adaptation. In interpreting the subgroup findings, caution is warranted. The extreme-obesity stratum (BMI ≥50 kg/m²) included only seven women, leading to imprecise estimates (wide CIs) that are sensitive to single observations. The study was not powered to detect differences within very high BMI classes; therefore these results should be viewed as hypothesis-generating and confirmed in larger cohorts.

Limitations

This single-center retrospective study may limit generalizability. To create balanced groups we randomly sampled 100 women per group from 2,600 eligible records; therefore, selection bias and limited representativeness are possible. Small numbers in some strata (e.g., BMI ≥50) reduce power, including for interaction tests. Birthweight centiles were unavailable, so SGA could not be assessed; to partially address confounding by gestational age we refit models for LBW, Apgar, and NICU including gestational age-LBW associations in PE(+) groups persisted, whereas Apgar attenuated and only PE(+) Ob(+) remained significant for NICU. Model diagnostics were acceptable overall, yet small cell counts warrant caution; multiple outcomes were tested without multiplicity adjustment, and GWG abstracted from routine records may contain measurement error. Finally, none of the participants received low-dose aspirin, which may limit external generalizability.

CONCLUSION

In our spline analysis, it was found that an increase in GWG during pregnancy in the PE(+) Ob(+) group was associated with a slight decrease in the risk of preterm birth and a partial improvement in Apgar scores. This result implies that putting too many limits on weight gain in the PE(+) Ob(+) group may be detrimental for perinatal health. Mechanistically, when a woman has preeclampsia and doesn't gain enough weight, it may further impair uteroplacental perfusion and limit the nourishment available to the fetus. However, when interpreting this trend in the spline curves, it should be kept

in mind that the relatively small sample size in the groups and the effect of outliers on the model may have influenced the results. Therefore, larger, prospective studies are needed to determine the optimal weight gain in the coexistence of preeclampsia and obesity.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Bursa City Hospital Scientific Researches Ethics Committee (Date: 25.06.2025, Decision No: 2025-13/8).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Availability of data and materials

The datasets used and analysed during the current study available from the corresponding author on reasonable request.

Acknowledgments

The author gratefully acknowledges anonymous support in statistical analysis and English language editing.

REFERENCES

- Poston L, Caleyachetty R, Cnattingius S, et al. Preconceptional and maternal obesity: epidemiology and health consequences. *Lancet Diabetes Endocrinol.* 2016;4(12):1025-1036. doi:10.1016/S2213-8587(16)30217-0
- Langley-Evans SC, Pearce J, Ellis S. Overweight, obesity and excessive weight gain in pregnancy as risk factors for adverse pregnancy outcomes: a narrative review. *J Hum Nutr Diet.* 2022;35(2):250-264. doi:10.1111/jhn.12999
- Sun Y, Shen Z, Zhan Y, et al. Effects of pre-pregnancy body-mass index and gestational weight gain on maternal and infant complications. *BMC Pregnancy Childbirth.* 2020;20(1):390. doi:10.1186/s12884-020-03071-y
- Zehravi M, Maqbool M, Ara I. Correlation between obesity, gestational diabetes mellitus, and pregnancy outcomes: an overview. *Int J Adolesc Med Health.* 2021;33(6):339-345. doi:10.1515/ijamh-2021-0058
- Marchi J, Berg M, Dencker A, Olander EK, Begley C. Risks associated with obesity in pregnancy, for the mother and baby: a systematic review of reviews. *Obes Rev.* 2015;16(8):621-638. doi:10.1111/obr.12288
- Crane JM, Murphy P, Burrage L, Hutchens D. Maternal and perinatal outcomes of extreme obesity in pregnancy. *J Obstet Gynaecol Can.* 2013;35(7):606-611. doi:10.1016/S1701-2163(15)30879-3
- Erez O, Romero R, Jung E, et al. Preeclampsia and eclampsia: the conceptual evolution of a syndrome. *Am J Obstet Gynecol.* 2022;226(2S):S786-S803. doi:10.1016/j.ajog.2021.12.001
- Suksai M, Geater A, Amornchat P, Suntharasing T, Suwanrath C, Pruksanusak N. Preeclampsia and timing of delivery: disease severity, maternal and perinatal outcomes. *Pregnancy Hypertens.* 2024;37:101151. doi:10.1016/j.preghy.2024.101151
- Poniedziałek-Czajkowska E, Mierzyński R, Leszczyńska-Gorzelak B. Preeclampsia and obesity-the preventive role of exercise. *Int J Environ Res Public Health.* 2023;20(2):1267. doi:10.3390/ijerph20021267
- Gestational hypertension and preeclampsia: ACOG practice bulletin, number 222. *Obstet Gynecol.* 2020;135(6):e237-e260. doi:10.1097/AOG.0000000000003891
- Rasmussen KM, Yaktine AL, Institute of Medicine (US) and National Research Council (US) Committee to Reexamine IOM Pregnancy Weight Guidelines, eds. *Weight Gain During Pregnancy: Reexamining the Guidelines.* Washington (DC): National Academies Press (US); 2009.
- Abalos E, Cuesta C, Carroli G, et al. Pre-eclampsia, eclampsia and adverse maternal and perinatal outcomes: a secondary analysis of the World Health Organization Multicountry Survey on Maternal and Newborn Health. *BJOG.* 2014;121(Suppl 1):14-24. doi:10.1111/1471-0528.12629
- Cnattingius S, Villamor E, Johansson S, et al. Maternal obesity and risk of preterm delivery. *JAMA.* 2013;309(22):2362-2370. doi:10.1001/jama.2013.6295
- Li L, Chen Y, Lin Z, et al. Association of pre-pregnancy body-mass index with adverse pregnancy outcome among first-time mothers. *PeerJ.* 2020;8:e10123. doi:10.7717/peerj.10123
- Su XJ, Huang SJ, Li X, Du QL. Prepregnancy overweight and obesity are associated with an increased risk of preterm birth in Chinese women. *Obes Facts.* 2020;13(2):237-244. doi:10.1159/000506688
- Jeyabalan A. Epidemiology of preeclampsia: impact of obesity. *Nutr Rev.* 2013;71(Suppl 1):S18-S25. doi:10.1111/nure.12055
- McDonald SD, Han Z, Mulla S, Beyene J; Knowledge Synthesis Group. Overweight and obesity in mothers and risk of preterm birth and low birth weight infants: systematic review and meta-analyses. *BMJ.* 2010;341:c3428. doi:10.1136/bmj.c3428
- Lewandowska M. Maternal obesity and risk of low birth weight, fetal growth restriction, and macrosomia: multiple analyses. *Nutrients.* 2021;13(4):1213. doi:10.3390/nu13041213
- Phipps E, Prasanna D, Brima W, Jim B. Preeclampsia: updates in pathogenesis, definitions, and guidelines. *Clin J Am Soc Nephrol.* 2016;11(6):1102-1113. doi:10.2215/CJN.12081115
- McIntyre HD, Gibbons KS, Flenady VJ, Callaway LK. Overweight and obesity in Australian mothers: epidemic or endemic? *Med J Aust.* 2012;196(3):184-188. doi:10.5694/mja11.11120
- Mbah AK, Kornosky JL, Kristensen S, et al. Super-obesity and risk for early and late pre-eclampsia. *BJOG.* 2010;117(8):997-1004. doi:10.1111/j.1471-0528.2010.02593.x
- O'Brien TE, Ray JG, Chan WS. Maternal body-mass index and the risk of preeclampsia: a systematic overview. *Epidemiology.* 2003;14(3):368-374. doi:10.1097/00001648-200305000-00020
- Melchor I, Burgos J, Del Campo A, Aiartzagüena A, Gutiérrez J, Melchor JC. Effect of maternal obesity on pregnancy outcomes in women delivering singleton babies: a historical cohort study. *J Perinat Med.* 2019;47(6):625-630. doi:10.1515/jpm-2019-0103
- Asemota O, Bryson L, Fogel J. Neonatal outcomes in obese pregnant women who develop preeclampsia. *Pregnancy Hypertens.* 2020;21:191-196. doi:10.1016/j.preghy.2020.05.016

Epidemiology of viral respiratory tract infections in pediatric ICU: a single-center study in Türkiye

 Abdulgani Gülyüz¹,  Nazan Poyraz²,  Sıla Yıldız Ataş²

¹Department of Child Health and Diseases, Faculty of Medicine, Malatya Turgut Özal University, Malatya, Türkiye

²Department of Child Health and Diseases, Malatya Training and Research Hospital, Malatya, Türkiye

Cite this article as: Gülyüz A, Poyraz N, Yıldız Ataş S. Epidemiology of viral respiratory tract infections in pediatric ICU: a single-center study in Türkiye. *J Health Sci Med.* 2025;8(6):1099-1104.

Received: 06.08.2025

Accepted: 10.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aims to evaluate the distribution of viral agents and their relationship with clinical parameters in pediatric patients with respiratory tract infections in the pediatric intensive care unit (PICU) of MALATYA Training and Research Hospital during 2024.

Methods: This retrospective observational study included 257 patients admitted to the PICU with a prediagnosis of respiratory tract infection between January 1 and December 31, 2024. Nasopharyngeal swab samples obtained from the patients were analyzed by multiplex PCR method. Demographic data, laboratory data and clinical course of the patients were recorded and statistical analysis was performed.

Results: The most common viral agents were rhinovirus/enterovirus (36.6%) and respiratory syncytial virus (RSV) (31.5%). The comorbidity rate was 38.1%. RSV infections were associated with significantly longer ICU stay compared to rhinovirus (mean 6.8 days vs. 4.3 days in the rhinovirus group, $p=0.008$). According to multiple regression analysis, age, CRP level and presence of comorbidities were determined as factors that significantly affected the length of stay ($p<0.05$).

Conclusion: Rhinovirus/enterovirus and RSV are the most common viral respiratory tract agents in PICU patients. RSV infection is associated with longer hospitalization. Elevated CRP and the presence of comorbid diseases also negatively affect the clinical course. These findings emphasize the importance of rapid diagnosis and consideration of risk factors in pediatric intensive care management.

Keywords: Pediatric intensive care, viral agents, rhinovirus, respiratory syncytial virus, comorbidity, PCR

INTRODUCTION

Respiratory tract infections are among the most common infectious diseases in childhood and may lead to significant morbidity and mortality, especially in infants.¹ They are one of the main causes of hospitalizations in pediatric intensive care units (PICUs).¹⁻³ The most frequent viral pathogens include rhinovirus, respiratory syncytial virus (RSV), and influenza viruses.³⁻⁵ Particularly, RSV is the leading cause of lower respiratory tract infections and bronchiolitis in infants and is closely associated with oxygen support and intensive care requirements.^{3,6}

In recent years, molecular diagnostic methods providing rapid and reliable results have become widely used in clinical practice. Multiplex PCR panels allow the simultaneous detection of multiple viral agents, guiding clinical management effectively.⁷⁻¹² However, understanding the epidemiologic characteristics and clinical impact of these infections is as important as rapid diagnosis for optimizing patient care and infection control strategies.¹³

RSV mainly causes bronchiolitis and pneumonia, with relatively low C-reactive protein (CRP) levels due to limited

systemic inflammation.^{11,14} In contrast, rhinovirus infections usually have a milder course but can sometimes lead to severe respiratory distress and prolonged hospitalization.¹⁵⁻¹⁷

Furthermore, comorbidities significantly influence the clinical course and extend PICU stay. Chronic conditions such as asthma, cerebral palsy, epilepsy, and congenital heart diseases increase susceptibility to infections and complicate treatment.¹⁸⁻²⁰

Therefore, this study retrospectively evaluated the distribution of viral respiratory pathogens and their relationship with clinical parameters among children hospitalized in the PICU of Malatya Training and Research Hospital.

METHODS

Ethical Approval

The study was approved by the Malatya University Non-interventional Clinical Researches Ethics Committee (Date: 31.01.2025, Decision No: 2025/20), which allowed the anonymous use of the data and removed the requirement for

Corresponding Author: Abdulgani Gülyüz, abdulganigulyuz@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

patient consent. All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Study Design and Participants

This study was conducted as a retrospective observational study on pediatric patients hospitalized in the pediatric intensive care unit (PICU) of Malatya Training and Research Hospital between January 1 and December 31, 2024. The study included 257 pediatric patients who were treated in the unit and admitted with a prediagnosis of respiratory tract infection. The age range of the patients included in the study was between 1 month and 18 years, and the demographic data, clinical findings, laboratory results, and treatment processes of the patients were evaluated retrospectively.

The PICU of Malatya Training and Research Hospital is a 6-bed tertiary-level unit, managed by an experienced pediatric specialist with advanced training in intensive care. Although the unit does not currently provide subspecialty training in pediatric intensive care, it is organized in accordance with national standards.

Inclusion and Exclusion Criteria

Inclusion criteria: (1) Patients hospitalized in the PICU of Malatya Training and Research Hospital between 2024. (2) Patients who were treated in the intensive care unit due to respiratory tract infection. (3) Patients who underwent multiplex PCR test after nasopharyngeal swab sampling.

Exclusion criteria: (1) Patients being treated in intensive care for reasons other than respiratory tract infection. (2) Patients with incomplete or invalid PCR test results. (3) Patients with incomplete clinical data or lack of data required for analysis.

Data Collection

The data used in the study were collected from patients' electronic patient files. Data included demographic information, clinical findings, respiratory tract agents, comorbidities, CRP, white blood cell (WBC) count and length of intensive care unit stay.

Identification of Respiratory Pathogens

Respiratory tract agents were analyzed by multiplex PCR method. In the tests performed using nasopharyngeal samples, viral agents such as coronavirus, parainfluenza virus, adenovirus, SARS COV-2, bocavirus, rhinovirus/enterovirus, respiratory syncytial virus, influenza A and B, human metapneumovirus were investigated. The multiplex PCR kit used could not distinguish between rhinovirus and enterovirus due to their homology. Therefore, they are mentioned together in this study. Negative PCR results were also obtained in some patients. Mycoplasma, chlamydia and bordetella bacteria were also tested in the same test. Additional tests for bacterial co-infections were not performed. The study focused on viral agents; bacterial agents were not evaluated. This limits the diagnostic scope of the study.

Statistical Analysis

IBM SPSS Statistics 25 program was used for statistical analysis of the data obtained in the study. Data were

evaluated with descriptive statistics and reported as numbers, percentages, mean and standard deviation. Kruskal-Wallis H test was used for intergroup comparisons. The relationships between the causative agents of respiratory tract infections and clinical parameters were analyzed in this way. A value of $p < 0.05$ was considered statistically significant.

The distribution of data was evaluated using the Kolmogorov-Smirnov and Shapiro-Wilk tests. Non-normally distributed data were expressed as median (IQR 25-75), while normally distributed variables were presented as mean $\pm$ standard deviation." The relevant tables have also been updated to reflect this format.

Multiple linear regression analysis was applied to determine the independent factors affecting the duration of intensive care unit stay. In this model, 'length of ICU stay (days)' was included as the dependent variable and age (months), gender, WBC, CRP level and presence of comorbidities were included as independent variables. Regression results are reported for each variable with beta coefficient (B), 95% confidence interval and p-value. Analyses were performed using SPSS Statistics 25. Age, CRP and presence of comorbidity significantly affected the length of hospitalization ($p < 0.05$).

RESULTS

A total of 257 patients met the study inclusion criteria after applying the exclusion criteria (patients admitted to the PICU for reasons other than respiratory tract infection, patients with incomplete or invalid PCR test results, and patients with insufficient clinical data were excluded). These patients had been hospitalized in the PICU of Malatya Training and Research Hospital between January 1 and December 31, 2024, with a diagnosis of respiratory tract infection. Their nasopharyngeal and/or tracheal aspirate samples were analyzed using the Multiplex PCR method.

The PICU of Malatya Training and Research Hospital is a 6-bed tertiary-level unit, managed by an experienced pediatric specialist with advanced intensive care training. Although the unit does not currently provide formal pediatric intensive care subspecialty education, it operates in accordance with national pediatric critical care standards. During the study period, a total of 499 patients were admitted to the PICU for various reasons, of whom 257 were followed for respiratory tract infections. Among these patients, 18 (7.0%) required invasive mechanical ventilation, 72 (28.0%) received noninvasive ventilation, and the remainder were supported with high-flow nasal cannula (HFNC) or oxygen mask therapy.

The median age of the patients was 39.5 months (IQR 14-78 months). Regarding gender distribution, 152 (59.1%) were male and 105 (40.9%) were female. According to age groups, 109 (42.4%) patients were aged 0-12 months, 85 (33.1%) were 13-60 months, and 63 (24.5%) were over 60 months.

The number of patients with comorbidity was 98/257 (38.1%). As a result of the chi-square correspondence test, it was found that gender distribution showed a statistically significant difference ($p = 0.034$) ([Table 1](#)).

Table 1. Demographic and clinical characteristics of patients hospitalized in the PICU due to respiratory tract infection

Variable	n (%) / mean $\pm$ SD (or median, IQR)
Total number of patients	257
Mean age (months $\pm$ SD)	39.5 $\pm$ 26.0
Median age (IQR 25-75, months)	38 (14-78)
Minimum age (months)	1
Maximum age (months)	214
Male	152 (59.1%)
Female	105 (40.9%)
Invasive mechanical ventilation	18 (7.0%)
Noninvasive mechanical ventilation	72 (28.0%)
HFNC/oxygen support	167 (65.0%)

PICU: Pediatric intensive care unit, SD: Standard deviation, IQR: Inter quantile range, HFNC: High-flow nasal cannula

Most Common Respiratory Infection Agents

A total of 94 (36.6%) patients had rhinovirus/enterovirus, 81 (31.5%) had RSV, 14 (5.4%) had influenza A+B, and 7 (2.7%) had adenovirus, SARS-CoV-2 was detected in 8 (3.1%) patients, human metapneumovirus A/B in 7 (2.7%) patients, bocavirus in 3 (1.2%) patients and parainfluenza in 10 (3.9%) patients.

In 47 (18.3%) patients who underwent PCR test, viral agent could not be detected and the test result was evaluated as negative. In addition, more than one viral agent was detected in 48 patients (18.7%). In 18 of these patients, RSV was found positive together with another agent. In addition, Bordetella and Mycoplasma results, which were positive in 8 patients in the same panel, were excluded (Table 2).

Table 2. Viral agent distribution of PCR results

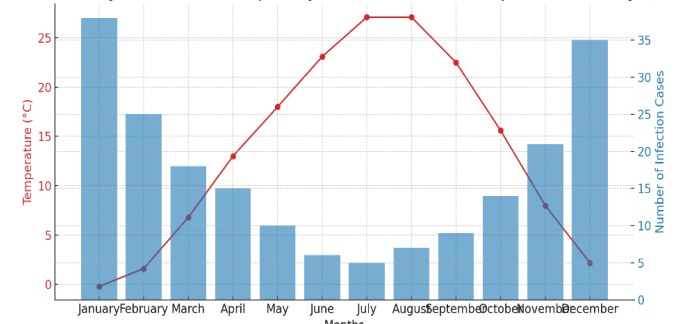
Viral agent	Number of cases (257)
Rhinovirus/enterovirus	94 (36.6%)
RSV	81 (31.5%)
Influenza a+b	14 (5.4%)
Parainfluenza	10 (3.9%)
Sarscov-2	8 (3.1%)
Adenovirus	7 (2.7%)
Human metapneumovirus	7 (2.7%)
Bocavirus	3 (1.1%)
Double factor (18 cases with RSV)	48
PCR negative	47

PCR: Polymerase chain reaction, RSV: Respiratory syncytial virus

Season and Temperature Effect

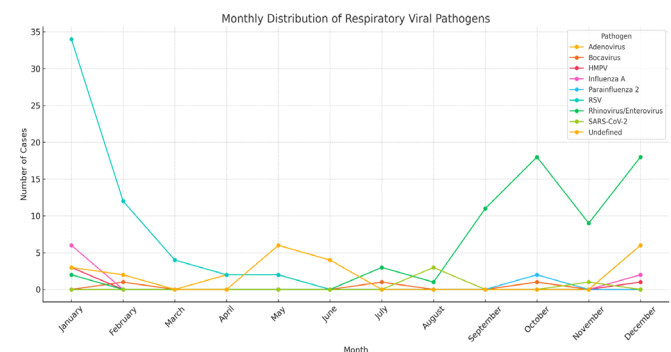
While the mean temperature in January was -0.2°C, this value was 28°C in June. An inverse relationship was observed between the number of monthly cases and the mean temperature values (Figure 1).

RSV cases increased significantly in December and January, while rhinovirus/enterovirus cases increased in December.

Figure 1. Comparison of number of infection cases and average air temperature**Figure 2. Monthly Distribution of Respiratory Infection Cases and Temperature in Malatya (20**

Influenza infections were also found to occur in December-January, coinciding with the winter season.

Adenovirus cases were common in May and December, suggesting that this agent may increase in both spring and winter (Figure 2).

**Figure 2. Distribution of respiratory tract infection agents according to months**

Intensive Care Unit Length of Stay and Acute Phase Reactants

There were no significant differences in CRP values between the diagnostic groups. The mean CRP value was 2.3 mg/L in the rhinovirus/enterovirus group and 2.5 mg/L in the RSV group. In patients with negative PCR results, the mean CRP value was 3.3 mg/L. No significant differences were found between WBC count and agent types. The mean WBC value was $12.2 \times 10^3/\mu\text{L}$ in the rhinovirus/enterovirus group and $11.5 \times 10^3/\mu\text{L}$ in the RSV group.

The duration of hospitalization in intensive care unit showed significant differences according to the type of agent. While patients with RSV infection were hospitalized for the longest period of time, patients in the rhinovirus/enterovirus group had shorter hospitalizations. The mean length of intensive care unit stay was 6.8 days in the RSV group and 4.3 days in the rhinovirus/enterovirus group. A shorter hospitalization period was observed in PCR-negative patients. The length of hospitalization was statistically significantly shorter in the PCR negative group compared to the RSV group ($p=0.002$). This indicates that the length of hospitalization may be slightly longer in patients with dual viral agents than in patients with a single agent, but the difference did not fully cross the threshold of statistical significance (Table 3).

Table 3. Mean CRP, WBC and length of hospitalization according to viral respiratory tract agents and p-values for pairwise comparisons

Group	Mean CRP (mg/L)	Mean WBC (x10 ³ /μL)	Average length of stay days	p with RSV	p with rhinovirus	p with double factor	PCR negative with p	CRP p-value	WBC p-value
RSV	2.5	11.5	6.8	-	0.008	0.217	0.002	-	-
Rhinovirus/enterovirus	2.3	12.2	4.3	0.008	-	0.091	0.076	0.284	0.325
Double agent	3.1	10.7	5.4	0.217	0.091	-	0.051	0.185	0.149
PCR negative	3.3	8.9	3.1	0.002	0.076	0.051	-	0.091	0.078

CRP: C-reactive protein, WBC: White blood cell, RSV: Respiratory syncytial virus, PCR: Polymerase chain reaction

According to multivariate regression analysis, age, CRP level and presence of comorbidity had a significant effect on the length of ICU stay. As age ($\beta=0.032$, $p=0.024$) and CRP level ($\beta=0.457$, $p=0.012$) increased, the length of hospitalization increased. In patients with comorbidities, the average length of hospitalization was 2.47 days longer ($\beta=2.470$, $p=0.001$). There was no significant relationship between gender ($p=0.641$) and WBC level ($p=0.428$) and length of hospitalization (Table 4).

Table 4. Multivariate regression analysis

Independent variable	β coefficient (B)	95% CI	p-value
Age (month)	0.032	0.011-0.053	0.024
CRP (mg/L)	0.457	0.10-0.81	0.012
WBC (x10 ³ /μL)	0.051	-0.08-0.18	0.428
Gender	0.273	-0.85-1.39	0.641
Comorbidity	2.47	1.21-3.73	0.001

CI: Confidence interval, CRP: C-reactive protein, WBC: White blood cell

Comorbidities

Comorbid diseases were found in 38.1% of the patients. The most common comorbid diseases were asthma and cerebral palsy (CP). Asthma was found in 13% and CP in 10% of patients followed up in intensive care unit. Comorbid diseases prolonged the duration of intensive care unit hospitalization and affected the treatment processes.

As a result of multivariate regression analysis, variables that had a significant effect on intensive care unit length of stay were determined. A slight increase in length of stay was observed with increasing age. An increase in CRP level also significantly prolonged the length of hospitalization. Intensive care unit length of stay was found to be 2.47 days longer on average in patients with comorbidities compared to those without. Gender and WBC levels had no significant effect on length of stay.

DISCUSSION

In this study, respiratory viral pathogens and related clinical parameters were retrospectively analyzed among 257 pediatric patients hospitalized with lower respiratory tract infections in the PICU of Malatya Training and Research Hospital during 2024.

Rhinovirus/enterovirus (36.6%) and RSV (31.5%) were the most frequently detected agents, consistent with multicenter

studies in the literature.^{11,21} Esen et al.⁸ reported RSV in 27.9% of hospitalized children, while Yüksel et al.²² found RSV and adenovirus at similar rates (31.5%). Pelit et al.²³ also identified RSV as the most common pathogen (35.3%).

SARS-CoV-2 positivity was 3.1%, reflecting post-pandemic shifts in viral distribution. Literature shows SARS-CoV-2 generally presents with milder illness and less need for intensive care than RSV and rhinovirus.²⁴ Since our data represent 2024, they provide updated epidemiologic insight for the post-COVID-19 period.

High detection rates of rhinovirus/enterovirus may be related to regional factors, patient age range, and sensitivity of molecular methods. Post-pandemic changes in viral circulation have been widely reported,^{5,21,25} suggesting altered pathogen dynamics globally.

Coinfection was found in 18.7% of patients, consistent with reported rates of 10-35%.^{5,25} RSV was the most frequent co-pathogen (n=18), supporting its known tendency to coexist with other viruses. However, the exact clinical impact of viral coinfections remains unclear and warrants further research.²⁶

Patients with rhinovirus/enterovirus infection had shorter ICU stays, likely due to milder disease limited to the upper airways, while RSV was associated with longer hospitalizations and more severe respiratory distress.^{14,27,28}

CRP and WBC were analyzed as markers of disease severity. Although differences were not statistically significant, the trends were clinically notable.²⁹ Low CRP levels in RSV infections support previous findings suggesting localized inflammation.^{14,29} WBC levels tended to be higher in rhinovirus cases and lower in RSV, though these alone are not reliable for viral differentiation.

Comorbidities were present in 38.1% of patients, most commonly asthma and cerebral palsy (CP), aligning with prior reports.³⁰ Chronic diseases increase infection susceptibility and hospitalization duration.^{18,19,31} Conditions such as prematurity, asthma, congenital heart disease, and neurodevelopmental disorders are recognized risk factors for severe respiratory illness requiring intensive care.^{20,32} Studies also show that comorbidity rates among PICU admissions are rising,² highlighting increasing clinical complexity.

In regression analysis, age, CRP level, and presence of comorbidities were independent predictors of longer PICU stay. This suggests that systemic inflammation and chronic

conditions contribute to prolonged hospitalization and greater clinical severity.

Seasonal distribution patterns were also consistent with previous research. RSV and rhinovirus infections peaked in winter months, particularly December-January.^{22,33,34} Cold temperatures, indoor crowding, and reduced immunity are likely contributing factors. These findings support increasing PICU capacity and infection control measures, including the use of palivizumab prophylaxis during high-risk months.³⁵

In conclusion, our findings align with national and international data and provide valuable post-pandemic insight into viral distribution, seasonality, and clinical outcomes in pediatric intensive care.

Limitations

Our findings are valuable as they include up-to-date data from the post-pandemic period. However, this study has some limitations:

- The single-center design limits the generalizability of the results. More comprehensive data can be obtained with multicenter studies from different regions.
- Only viral agents were focused and bacterial co-infections were excluded. This may lead to failure to detect mixed infections and incomplete interpretation of the clinical course.
- The timing of sampling is not standardized, which can affect the sensitivity of PCR tests, especially in some cases, leading to false negative results.

Taking these methodological limitations into account in future studies will both increase diagnostic accuracy and strengthen the reliability of the data obtained.

CONCLUSION

The most common viral agents in our pediatric ICU are rhinovirus/enterovirus and RSV, and RSV infections have a more severe course. High CRP levels and presence of comorbidities were associated with longer hospitalization. These findings emphasize the importance of rapid diagnosis and precautions for risk groups.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Malatya University Non-interventional Clinical Researches Ethics Committee (Date: 31.01.2025, Decision No: 2025/20).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Troeger C, Blacker B, Khalil IA, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis*. 2018;18(11):1191-1210.
2. Killien EY, Keller MR, Watson RS, Hartman ME. Epidemiology of intensive care admissions for children in the US from 2001 to 2019. *JAMA Pediatr*. 2023;177(5):506-515. doi:10.1001/jamapediatrics.2023.0184
3. Li Y, Wang X, Blau DM, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in children younger than 5 years in 2019: a systematic analysis. *Lancet*. 2022;399(10340):2047-2064. doi:10.1016/S0140-6736(22)00478-0
4. Perez A, Lively JY, Curns A, et al. Respiratory virus surveillance among children with acute respiratory illnesses-new vaccine surveillance network, United States, 2016-2021. *Morb Mortal Wkly Rep*. 2022;71(40):1253-1259.
5. Meissner HC. Viral bronchiolitis in children. *N Engl J Med*. 2016;374(1):62-72. doi:10.1056/NEJMra1413456
6. Hall CB, Weinberg GA, Iwane MK, et al. The burden of respiratory syncytial virus infection in young children. *N Engl J Med*. 2009;360(6):588-598. doi:10.1056/NEJMoa0804877
7. Brendish NJ, Malachira AK, Armstrong L, et al. Routine molecular point-of-care testing for respiratory viruses in adults presenting to hospital with acute respiratory illness (ResPOC): a pragmatic, open-label, randomized controlled trial. *Lancet Respir Med*. 2017;5(5):401-411. doi:10.1016/S2213-2600(17)30120-0
8. Esen AB, Erol M, Kafadar D, et al. Progression of disease and viral agents in infants hospitalized for lower respiratory tract infections. *J Pediatr Res*. 2019;6(4):314-321. doi:10.4274/jpr.galenos.2019.25348
9. Mahony JB. Detection of respiratory viruses by molecular methods. *Clin Microbiol Rev*. 2008;21(4):716-747. doi:10.1128/cmr.00037-07
10. Lin CY, Hwang D, Chiu NC, et al. Increased detection of viruses in children with respiratory tract infection using PCR. *Int J Environment Res Public Health*. 2020;17(2):564. doi: 10.3390/ijerph17020564
11. Esposito S, Mencacci A, Cenci E, Camilloni B, Silvestri E, Principi N. Multiplex platforms for the identification of respiratory pathogens: are they useful in pediatric clinical practice? *Front Cell Infect Microbiol*. 2019;9:196. doi:10.3389/fcimb.2019.00196
12. Clark TW, Lindsley K, Wigmosta TB, et al. Rapid multiplex PCR for respiratory viruses reduces time to result and improves clinical care: results of a systematic review and meta-analysis. *J Infect*. 2023;86(5):462-475. doi:10.1016/j.jinf.2023.03.005
13. Ljubin-Sternak S, Meštrović T. Rhinovirus-A true respiratory threat or a common inconvenience of childhood? *Viruses*. 2023;15(4):825. doi:10.3390/v15040825
14. Calvo C, García-García ML, Blanco C, et al. Multiple simultaneous viral infections in infants with acute respiratory tract infections in Spain. *J Clin Virol*. 2008;42(3):268-272. doi:10.1016/j.jcv.2008.03.012
15. Heikkinen T, Järvinen A. The common cold. *Lancet*. 2003;361(9351):51-59. doi:10.1016/S0140-6736(03)12162-9
16. Pneumonia Etiology Research for Child Health (PERCH) Study Group. Causes of severe pneumonia requiring hospital admission in children without HIV infection from Africa and Asia: the PERCH multi-country case-control study. *Lancet*. 2019;394(10200):757-779. doi:10.1016/S0140-6736(19)30721-4
17. Jacobs SE, Lamson DM, St George K, Walsh TJ. Human rhinoviruses. *Clin Microbiol Rev*. 2013;26(1):135-162. doi:10.1128/CMR.00077-12

18. Ghazaly M, Nadel S. Characteristics of children admitted to intensive care with acute bronchiolitis. *Eur J Pediatr*. 2018;177(6):913-920. doi:10.1007/s00431-018-3138-6
19. Havdal LB, Boås H, Bekkevold T, et al. Risk factors associated with severe disease in respiratory syncytial virus infected children under 5 years of age. *Front Pediatr*. 2022;10:1004739. doi:10.3389/fped.2022.1004739
20. Kirolos N, Mtaweh H, Datta RR, et al. Risk factors for severe disease among children hospitalized with respiratory syncytial virus. *JAMA Netw Open*. 2025;8(4):e254666. doi:10.1001/jamanetworkopen.2025.4666
21. Lei C, Yang L, Lou CT, et al. Viral etiology and epidemiology of pediatric patients hospitalized for acute respiratory tract infections in Macao: a retrospective study from 2014 to 2017. *BMC Infect Dis*. 2021;21(1):1-11. doi:10.1186/s12879-021-05996-x
22. Yüksel H, Yılmaz O, Akçali S, et al. Common viral etiologies of community acquired lower respiratory tract infections in young children and their relationship with long term complications. *Mikrobiyol Bul*. 2008;42(3):429-435.
23. Pelit S, Bayraktar B, Bulut ME, Karabulut ND, Nuhoglu A. Investigation of viral agents in lower respiratory tract infections of children by immunofluorescent and immunochromatographic methods. *Şişli Etfal Hastanesi Tıp Bülteni*, 2015;49(2):118.
24. Khales P, Razizadeh MH, Ghorbani S, Moattari A, Saadati H, Tavakoli A. Prevalence of respiratory viruses in children with respiratory tract infections during the COVID-19 pandemic era: a systematic review and meta-analysis. *BMC Pulm Med*. 2025;25(1):135. doi:10.1186/s12890-025-03587-z
25. Şık G, Demirbuğa A, Annayev A, Cabiri A, Deliceo E, Çıtak A. Frequency of viral pathogens and clinical characteristics of patients hospitalized with lower respiratory tract infection in pediatric intensive care unit. *J Pediatr Infect*. 2020;14(1):27-32. doi:10.5578/ced.202011
26. Scotta MC, Chakr VC, de Moura A, et al. Respiratory viral coinfection and disease severity in children: a systematic review and meta-analysis. *J Clin Virol*. 2016;80:45-56. doi:10.1016/j.jcv.2016.04.019
27. Zumla A. Killer respiratory tract infections: time to turn the tide. *Curr Opin Pulm Med*. 2012;18(3):173-174. Available from: doi:10.1097/MCP.0b013e328351f7d8
28. Jain S, Williams DJ, Arnold SR, et al. Community-acquired pneumonia requiring hospitalization among U.S. children. *N Engl J Med*. 2015; 372(9):835-845. doi:10.1056/NEJMoa1405870
29. Tregoning JS, Schwarze J. Respiratory viral infections in infants: causes, clinical symptoms, virology, and immunology. *Clin Microbiol Rev*. 2010; 23(1):74-98. doi:10.1128/cmr.00032-09?download=true
30. Berry JG, Hall M, Hall DE, et al. Inpatient growth and resource use in 28 children's hospitals: a longitudinal, multi-institutional study. *JAMA Pediatr*. 2013;167(2):170-177. doi:10.1001/jamapediatrics.2013.432
31. Krishnamurthy K, King R, Oredein B, et al. Admission trends, associated factors and the outcomes for children hospitalised to paediatric intensive care unit for asthma-a population based longitudinal study. *Karnataka Paediatr J*. 2025;39(4):125-129. doi:10.25259/KPJ_22_2024
32. Hatipoğlu S, Arica S, Çelik Y, et al. Frequency and clinical characteristics of RSV infection in patients hospitalized with lower respiratory tract infection. *Duzce Med J*. 2009;11(1):38-44.
33. Kudo E, Song E, Yockey LJ, et al. Low ambient humidity impairs barrier function and innate resistance against influenza infection. *Proc Natl Acad Sci U S A*. 2019;166(22):10905-10910. doi:10.1073/pnas.1902840116?download=true
34. Griffiths C, Drews SJ, Marchant DJ. Respiratory syncytial virus: infection, detection, and new options for prevention and treatment. *Clin Microbiol Rev*. 2017;30(1):277-319. doi:10.1128/CMR.00010-16
35. Hammitt LL, Dagan R, Yuan Y, et al. Nirsevimab for prevention of RSV in healthy late-preterm and term infants. *N Engl J Med*. 2022;386(9):837-846. doi:10.1056/NEJMoa2110275

How reliable are YouTube videos on dental implants? A comparison between Turkish and English content

 Ezgi Yüceer Çetiner¹,  Reşat Batuhan Çetiner²

¹Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Bahçeşehir University, İstanbul, Türkiye

²Department of Prosthodontics, Faculty of Dentistry, İstanbul Aydın University, İstanbul, Türkiye

Cite this article as: Yüceer Çetiner E, Çetiner RB. How reliable are YouTube videos on dental implants? A comparison between Turkish and English content. *J Health Sci Med.* 2025;8(6):1105-1111.

Received: 18.08.2025

Accepted: 10.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to evaluate and compare the reliability, online visibility, and educational quality of Turkish and English YouTube videos related to dental implants using validated scoring tools.

Methods: A cross-sectional study was conducted on YouTube videos retrieved on July 30, 2025, using the keywords “dental implant” (English) and “diş implantı” (Turkish). The first 100 videos in each language were screened, and 50 eligible videos per group were included after applying predefined inclusion and exclusion criteria. Videos were assessed for uploader profile, target audience, video duration, engagement metrics (views, likes, dislikes, comments, Video Power Index), and content quality using the DISCERN (short-form) and global quality score (GQS). Statistical analyses were performed using nonparametric tests and correlation analyses (Spearman's rho) with a significance level set at $p < 0.05$.

Results: English videos had significantly longer duration (median 262.5 vs. 121 minutes), higher VPI (69.26 vs. 2.93), more views (68,832.5 vs. 5,220), and more likes (665.5 vs. 18.5) compared with Turkish videos (all $p < 0.001$). DISCERN and GQS scores were significantly higher for English videos (median 3 vs. 2; $p = 0.007$ for both). While English videos demonstrated consistent positive correlations between quality scores and engagement metrics, Turkish videos showed no such pattern, and a negative correlation was observed between views and the like/view ratio ($r = -0.532$; $p < 0.001$).

Conclusion: Although English-language YouTube videos on dental implants demonstrated moderately higher reliability, quality, and engagement compared to Turkish videos, the overall educational value of both groups remained suboptimal. These findings highlight the urgent need for structured, evidence-based, and multilingual digital resources to ensure accurate and accessible patient education.

Keywords: Dental implants, online health information, patient education, Video Power Index, YouTube

INTRODUCTION

Dental implants have become a widely accepted and frequently performed treatment modality for the replacement of missing teeth.^{1,2} Offering superior functional outcomes, improved esthetics, and long-term durability compared to traditional prostheses, implants play a critical role in restoring oral health and quality of life.³ As implantology becomes increasingly common, the number of patients considering or undergoing this treatment has risen substantially worldwide. However, despite their widespread use, dental implants remain a complex and often anxiety-inducing topic for patients due to factors such as surgical procedures, healing timelines, costs, potential complications, and long-term maintenance.^{4,5} In this context, many patients turn to the internet as a primary source of information to better understand the procedure, its risks, benefits, and alternatives before making treatment decisions.^{6,7}

In recent years, YouTube has emerged as one of the most popular platforms for patients seeking health-related information due

to its ease of access, visual format, and perceived authenticity.^{8,9} However, the unregulated nature of user-generated content raises questions regarding the accuracy and reliability of the information presented.^{10,11} Several studies have investigated the quality of YouTube videos in various fields of dentistry and medicine, often revealing considerable variability in the credibility and educational value of the content.^{6,9,11-15}

Recent research has expanded the evaluation of YouTube as a platform for dental education. Tayan et al.¹⁶ analyzed dental implant videos and found that most were produced by healthcare professionals, while those incorporating patient experiences attracted higher levels of audience engagement. Similarly, Di Spirito et al.¹⁷ investigated peri-implantitis-related content and reported that the majority of videos had low educational value. Although Higher Quality Index and content scores were associated with better educational potential. Complementing these findings, Kuka and Gursoy¹⁸ evaluated oral hygiene videos and showed that only a minority

Corresponding Author: Ezgi Yüceer Çetiner, ezgi.yuceercetiner@bau.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

were of excellent quality, with longer duration and higher content scores contributing to better ratings. Together, these findings highlight the considerable variability in online dental resources and emphasize the need to assess video quality across different topics and languages to support accurate patient education. This variability also underscores the importance of using standardized instruments for objective evaluation.

Tools such as the DISCERN instrument and the global quality score (GQS) have been validated and frequently used to assess the reliability and overall quality of online health information.¹⁹ Despite the growing body of research on English-language content, there remains a paucity of comparative studies examining the quality of videos across different languages, particularly in non-English-speaking populations.²⁰

Language is a critical determinant of health communication, shaping both the accessibility and the trustworthiness of online resources. Patients often prefer to consume information in their native language, which may enhance understanding but does not guarantee accuracy or completeness. Prior multilingual evaluations of online health content have demonstrated notable discrepancies between languages. For instance, Berland et al.²¹ reported differences in the accessibility and quality of English versus Spanish health information on the internet, while a recent comparison of Turkish- and English-language YouTube videos on gingival recession also revealed lower quality scores in Turkish content.²⁰ These findings suggest that linguistic context may influence both the quantity and quality of information available to patients. Given the increasing number of Turkish-speaking patients seeking implant treatments both domestically and abroad, evaluating the quality of Turkish-language video content is of particular importance.

In light of these findings, the present study aims to compare the reliability, popularity, and overall quality of Turkish and English YouTube videos on dental implants using validated assessment tools. By identifying potential differences between the two language groups, this research seeks to highlight gaps in patient education and underscore the need for improved digital health communication strategies in diverse linguistic contexts.

METHODS

As this research did not involve human participants or human-derived samples and was carried out using publicly accessible content from a website, approval from an ethics committee was not necessary. Institutional permission was obtained from Bahçeşehir University Faculty of Dentistry to conduct and report the research. All procedures were carried out in accordance with the ethical rules and the principles.

Search Strategy

A YouTube search was performed on 30 July 2025 using the keywords “dental implant” (for English videos) and “diş implantı” (for Turkish videos). The relevance sorting filter was applied in both searches. To minimize algorithmic bias and personalization, all searches were conducted in incognito mode after clearing browsing history and cookies. The first

100 videos retrieved in each language were recorded in separate playlists. At the time of retrieval, metadata including the number of likes, dislikes, comments, video duration, and view counts was documented. In the following days, the videos preserved in these playlists were systematically reviewed for content evaluation.

Selection of videos: From the 100 videos screened per language, 50 eligible videos were included for detailed analysis. No additional computer-based randomization procedure was applied. Instead, we acknowledged the inherent variability of YouTube’s search algorithm as a natural element of randomness, since search results may differ depending on the time of query, location, and user history. By conducting the searches in incognito mode and saving playlists on the day of retrieval, we ensured that the same set of videos remained consistently available for subsequent evaluation.

Inclusion and exclusion criteria: Videos were included if they were directly related to dental implants, publicly accessible, non-sponsored, and accompanied by either audio narration or textual explanation, with a minimum resolution of 240p. Exclusion criteria applied to videos that were duplicates, irrelevant to the topic, purely promotional or advertisements, and those lacking sound or written explanation. In addition, videos shorter than 30 seconds or longer than 15 minutes. Consistent with earlier research, a 15-minute upper limit was applied, since prolonged videos generally fail to maintain audience attention.⁶ Also, videos with disabled comment sections or hidden “like” and “dislike” counts at the time of data collection, were excluded from the analysis to avoid bias in engagement-related analyses.

Data collection and assessment tools: For each video, descriptive parameters such as video duration (in seconds), upload date (day/month/year), total number of views, likes, dislikes, and comments were recorded. Viewer engagement was further quantified using the Video Power Index (VPI), calculated with the formula: $VPI = [(view\ rate \times like\ rate) / 100]$ where view rate is calculated as: (total views/days since upload), while like rate was calculated as $[(100 \times \text{number of likes}) / (\text{number of likes} + \text{number of dislikes})]$.²² In this study, the number of dislikes on YouTube videos was retrieved using the “YouTube Dislike Extension,” a Chrome browser add-on available through the Chrome Web Store, given that YouTube no longer publicly displays this data.²²

Each video was evaluated jointly by an oral and maxillofacial surgeon and a prosthodontist. The videos were viewed simultaneously at normal playback speed, and all evaluations were carried out by consensus. In cases of disagreement regarding scoring or classification, decisions were made through discussion and by consulting relevant literature. Inter-rater reliability between the two reviewers was assessed using the intraclass correlation coefficient (ICC) to ensure consistency in scoring. To ensure a standardized and reproducible assessment, two validated evaluation tools were used. The first, the GQS, assessed the overall quality, accuracy, and comprehensibility of each video from the patient’s perspective, using a five-point Likert Scale (1=poor, 5=excellent) (Table 1).²² The second tool, the short-form DISCERN instrument, consisted of five structured questions

aimed at evaluating the reliability, objectivity, and clarity of treatment options. Each item was scored as 1 for “yes” and 0 for “no”, yielding a total reliability score ranging from 0 to 5, with higher scores reflecting greater reliability of the video content (**Table 2**).¹⁹

Table 1. Global Quality Score Five Point Scale and score description

Five-Point Scale	Score description
1	Poor quality, poor flow of video, most information missing, not at all useful for patients
2	Generally poor quality and poor flow, some information listed but many important topics missing, of very limited use to patients
3	Moderate quality, suboptimal flow, some important information is adequately discussed but others poorly discussed, somewhat useful for patients
4	Good quality and generally flow. Most of the relevant information is listed, but some topics not covered, useful for patients
5	Excellent quality and flow, very useful for patients

Table 2. Reliability score (DISCERN-short form)

1. Are the aims clear and achieved?
2. Are reliable sources of information used?
3. Is the information presented balanced and unbiased?
4. Are additional sources of information listed for patient reference?
5. Are areas of uncertainty mentioned?

To further characterize the content, each video was categorized according to the identity of the uploader and the intended target audience. The uploader profile was classified into five groups: (1) dentist or dental clinic, (2) other healthcare professionals, (3) patient or blogger, (4) commercial company, and (5) media outlet (e.g., news organizations). The target audience was also classified into three categories: (1) patients, (2) dentists for educational purposes, and (3) dentists for commercial purposes.

Power analysis: An a priori power analysis was performed to estimate the minimum sample size required to detect a significant difference between groups. The analysis was conducted with G*power V3.1 software (Heinrich-Heine-Universität Düsseldorf, Düsseldorf, Germany) using the Wilcoxon-Mann-Whitney test (two-tailed, A.R.E. method). The input values were based on the video duration parameter reported by Yelbay et al.²⁰ The calculations assumed an effect size ($d=0.89$), significance level ($\alpha=0.05$), and desired power ($1-\beta=0.95$) with equal allocation ($N2/N1=1$). The required sample size was determined as 40 videos per group (80 in total). This confirmed that the present study's sample size was sufficient to achieve adequate statistical power.

Statistical Analysis

Data were analyzed using IBM SPSS Statistics, version 23 (IBM Corp., Armonk, NY, USA). The Kolmogorov-Smirnov test was applied to assess the normality of data distribution. For comparisons between two independent groups, the Mann-Whitney U test was used when variables did not follow a normal distribution. Associations between categorical

variables were examined with Fisher's exact test using the Monte Carlo correction, and pairwise comparisons were conducted with the Bonferroni-adjusted Z test. Interobserver agreement was evaluated with the ICC. Correlations between non-normally distributed quantitative variables were assessed using Spearman's rho correlation coefficient, and the significance threshold in correlation analyses was adjusted by the Bonferroni correction. Results were presented as frequencies and percentages for categorical variables, and as medians with minimum and maximum values for quantitative variables. A p-value <0.050 was considered statistically significant.

RESULTS

For the Turkish videos, interobserver agreement was statistically good for the DISCERN Scale ($ICC=0.862$; $p<0.001$) and excellent for the GQS Scale ($ICC=0.960$; $p<0.001$). For the English videos, interobserver agreement was statistically excellent for both the DISCERN ($ICC=0.950$; $p<0.001$) and GQS ($ICC=0.941$; $p<0.001$) scales. Overall, interobserver agreement was excellent for the DISCERN ($ICC=0.927$; $p<0.001$) and GQS ($ICC=0.951$; $p<0.001$) scales (**Table 3**).

Table 3. Interobserver reliability assessment

	Turkish	English	Total
DISCERN	0.862 (0.77-0.92)/ <0.001	0.95 (0.913-0.971)/ <0.001	0.927 (0.894-0.951)/ <0.001
GQS	0.96 (0.93-0.977)/ <0.001	0.941 (0.899-0.966)/ <0.001	0.951 (0.928-0.967)/ <0.001
Intraclass correlation coefficient (95% confidence interval)/p; intraclass correlation coefficient (95% confidence interval), GQS: Global quality score			

The median video duration was 262.5 minutes for English videos and 121 minutes for Turkish videos, with a statistically significant difference ($p<0.001$). The median VPI was 69.26 for English and 2.93 for Turkish videos ($p<0.001$). The median number of views was 68,832.5 for English and 5,220 for Turkish videos ($p<0.001$). The median number of likes was 665.5 for English and 18.5 for Turkish videos ($p<0.001$). The median number of dislikes was 9 for English and 0 for Turkish videos ($p<0.001$). The median number of comments was 93 for English and 0.5 for Turkish videos ($p<0.001$). The median number of days online was 1,093.5 for English and 1,269 for Turkish videos, with no significant difference ($p=0.617$). The median DISCERN score was 3 for English and 2 for Turkish videos ($p=0.007$). The median GQS score was also 3 for English and 2 for Turkish videos ($p=0.007$). The median like/view ratio was 0.01 for English and 0 for Turkish videos ($p=0.002$). A significant association was found between video language and content creator profile ($p=0.006$). Among English videos, 84% were uploaded by dentists or dental clinics, compared with 100% in Turkish videos (**Table 4**).

Regarding GQS score distribution, Turkish videos were most frequently rated as generally poor quality (56.0%), followed by moderate quality (22.0%), poor quality (10.0%), good quality (10.0%), and excellent quality (2.0%). In contrast, English videos more often achieved moderate quality ratings (36.0%), followed by generally poor quality (32.0%), good quality (20.0%), poor quality (6.0%), and excellent quality (6.0%).

Table 4. Comparison of variables by video language

	English	Turkish	Total	Test statistics	p	Effect size
Video duration	262.5 (58-900)	121 (35-811)	168 (35-900)	616.000	<0.001 ^x	0.972 [*]
VPI	69.26 (1.66-1430.33)	2.93 (0-110.4)	15.26 (0-1430.33)	375.000	<0.001 ^x	1.513 [*]
Number of views	68,832.5 (560-1,878,496)	5220 (73-267,189)	23,079.5 (73-1,878,496)	532.000	<0.001 ^x	1.139 [*]
Number of likes	665.5 (3-21,000)	18.5 (0-842)	100 (0-21,000)	310.000	<0.001 ^x	1.702 [*]
Number of dislikes	9 (0-1400)	0 (0-129)	0 (0-1400)	540.500	<0.001 ^x	1.122 [*]
Number of comments	93 (0-2649)	0.5 (0-477)	9.5 (0-2649)	479.500	<0.001 ^x	1.254 [*]
Number of days online	1093.5 (24-4558)	1269 (37-4488)	1115 (24-4558)	1177.500	0.617 [*]	0.100 [*]
DISCERN	3 (1-5)	2 (1-3)	2 (1-5)	877.500	0.007 [*]	0.531 [*]
GQS	3 (1-5)	2 (1-5)	2 (1-5)	882.000	0.007 [*]	0.525 [*]
Like/view ratio	0.01 (0-0.22)	0 (0-0.56)	0.01 (0-0.56)	810.000	0.002 [*]	0.637 [*]
Uploader profile						
Dentist/dental clinic	42 (84) ^a	50 (100) ^b	92 (92)	---	0.006 ^y	0.295 ^{**}
Other healthcare professionals	2 (4)	0 (0)	2 (2)			
Patient/blogger	3 (6)	0 (0)	3 (3)			
Commercial company	2 (4)	0 (0)	2 (2)			
Media outlet	1 (2)	0 (0)	1 (1)			
Target audience						
Patient	50 (100)	50 (100)	100 (100)	---	---	---
*Mann-Whitney U test; ^y Fisher's exact test with Monte Carlo correction; ^x Cohen's d; ^{**} Cramer's V; median (minimum-maximum); Frequency (percentage); ^{a-b} : No significant difference between groups sharing the same letter. VPI: Video Power Index, GQS: Global quality score						

*Mann-Whitney U test; ^yFisher's exact test with Monte Carlo correction; ^xCohen's d; ^{**}Cramer's V; median (minimum-maximum); Frequency (percentage); ^{a,b} No significant difference between groups sharing the same letter.

VPI: Video Power Index, GQS: Global quality score

When correlations among parameters were examined for the English videos, VPI showed very strong positive associations with views ($r=0.841$; $p<0.001$) and strong associations with likes ($r=0.773$; $p<0.001$) and dislikes ($r=0.620$; $p<0.001$). Views were strongly correlated with likes ($r=0.709$; $p<0.001$) and very strongly with dislikes ($r=0.801$; $p<0.001$), while likes and dislikes were also strongly correlated ($r=0.719$; $p<0.001$). Video duration showed a moderate correlation with comments ($r=0.492$; $p<0.001$). Comments were strongly associated with VPI ($r=0.730$; $p<0.001$), views ($r=0.755$; $p<0.001$), likes ($r=0.824$; $p<0.001$), and dislikes ($r=0.750$; $p<0.001$). A very strong correlation was observed between the DISCERN and GQS scores ($r=0.837$; $p<0.001$). VPI also showed a moderate correlation with the content creator profile ($r=0.507$; $p<0.001$). No other significant associations were identified (Table 5).

For the Turkish videos, VPI was strongly correlated with views ($r=0.664$; $p<0.001$) and very strongly with likes ($r=0.849$; $p<0.001$). Views were also strongly correlated with likes ($r=0.654$; $p<0.001$). VPI showed moderate correlations with dislikes ($r=0.477$; $p<0.001$), as did views ($r=0.517$; $p<0.001$) and likes ($r=0.593$; $p<0.001$). Views were strongly correlated with comments ($r=0.702$; $p<0.001$), while likes ($r=0.484$; $p<0.001$) and dislikes ($r=0.547$; $p<0.001$) showed moderate correlations with comments. A strong correlation was also observed between DISCERN and GQS scores ($r=0.658$; $p<0.001$). In contrast, views and the like/view ratio were moderately and negatively correlated ($r=-0.532$; $p<0.001$).

No other significant associations were detected. It should be noted that, because the target audience was limited to patients and all Turkish videos were uploaded exclusively by dentists or dental clinics, the variable "content creator profile" was not included in the correlation analysis (Table 6).

DISCUSSION

Dental implant procedures have gained prominence as a primary solution for restoring missing teeth, increasingly favored over removable prostheses by both clinicians and patients.² This rise in preference is not solely due to clinical success rates, but also reflects the growing influence of digital media and online platforms on treatment awareness.⁵ Despite the increasing number of videos on dental implants, there is limited information about their educational quality, especially in languages other than English. This study aimed to evaluate and compare the reliability and quality of Turkish and English YouTube videos on dental implants, with particular focus on uploader credentials, audience targeting, video popularity metrics, and content quality as assessed by DISCERN and GQS scores. Unlike most previous studies that focus only on English or only on Turkish videos, this study provides a comparative evaluation of both, revealing potential disparities in information quality across languages.^{4,15,16,25,26}

In the present study, English-language YouTube videos on dental implants demonstrated a moderate level of content quality, with a median DISCERN score and GQS of 3. These values suggest that while the videos provided some useful information, they often lacked comprehensive, evidence-based explanations or balanced discussions of treatment options. This finding is important, as it indicates that English videos, despite outperforming Turkish ones, cannot be regarded as wholly reliable educational resources. Their scores still fall into the moderate range, consistent with broader literature where most medical and dental videos on YouTube achieve only poor-to-moderate quality ratings.¹⁵⁻¹⁸ Together with these, Turkish-language videos showed lower performance

Table 5. Analysis of relationships between variables in English videos

		1	2	3	4	5	6	7	8	9	10	11
1. Video duration	r	1										
	p	.										
2. VPI	r	0.474	1									
	p	0.001	.									
3. Number of views	r	0.276	0.841	1								
	p	0.052	<0.001	.								
4. Number of likes	r	0.461	0.773	0.709	1							
	p	0.001	<0.001	<0.001	.							
5. Number of dislikes	r	0.376	0.620	0.801	0.719	1						
	p	0.007	<0.001	<0.001	<0.001	.						
6. Number of comments	r	0.492	0.730	0.755	0.824	0.750	1					
	p	<0.001	<0.001	<0.001	<0.001	<0.001	.					
7. Number of days online	r	-0.290	-0.150	0.331	0.043	0.361	0.135	1				
	p	0.041	0.299	0.019	0.766	0.010	0.350	.				
8. DISCERN	r	0.293	0.238	0.211	0.092	0.120	0.240	0.050	1			
	p	0.039	0.096	0.141	0.524	0.405	0.093	0.730	.			
9. GQS	r	0.276	0.200	0.241	0.117	0.164	0.241	0.178	0.837	1		
	p	0.052	0.163	0.092	0.417	0.256	0.092	0.216	<0.001	.		
10. Uploader profile	r	0.156	0.507	0.459	0.408	0.407	0.424	0.127	0.224	0.331	1	
	p	0.280	<0.001	0.001	0.003	0.003	0.002	0.378	0.118	0.019	.	
11. Like/view ratio	r	0.472	-0.022	-0.314	0.352	-0.055	0.137	-0.438	0.011	-0.042	-0.120	1
	p	0.001	0.882	0.026	0.012	0.705	0.344	0.001	0.940	0.772	0.405	.

r: Spearman's rho correlation coefficient. The numbers in the first row correspond to the variables listed in the vertical columns, enabling pairwise correlations to be presented both horizontally and vertically for the same set of variables. VPI: Video Power Index, GQS: Global quality score

Table 6. Analysis of relationships between variables in Turkish videos

		1	2	3	4	5	6	7	8	9	10
1. Video duration	r	1									
	p	.									
2. VPI	r	0.168	1								
	p	0.243	.								
3. Number of views	r	0.178	0.664	1							
	p	0.215	<0.001	.							
4. Number of likes	r	0.289	0.849	0.654	1						
	p	0.042	<0.001	<0.001	.						
5. Number of dislikes	r	0.277	0.477	0.517	0.593	1					
	p	0.052	<0.001	<0.001	<0.001	.					
6. Number of comments	r	0.258	0.436	0.702	0.484	0.547	1				
	p	0.070	0.002	<0.001	<0.001	<0.001	.				
7. Number of days online	r	-0.144	0.013	0.439	0.271	0.352	0.257	1			
	p	0.319	0.928	0.001	0.057	0.012	0.071	.			
8. DISCERN	r	0.544	0.031	0.107	0.107	0.220	0.350	-0.042	1		
	p	<0.001	0.831	0.459	0.459	0.125	0.013	0.772	.		
9. GQS	r	0.550	0.060	0.029	0.082	0.049	0.186	-0.261	0.658	1	
	p	<0.001	0.678	0.840	0.572	0.736	0.195	0.068	<0.001	.	
10. Like/view ratio	r	0.387	-0.073	-0.532	0.093	-0.029	-0.269	-0.365	0.269	0.281	1
	p	0.005	0.615	<0.001	0.523	0.844	0.059	0.009	0.059	0.048	.

r: Spearman's rho correlation coefficient. The numbers in the first row correspond to the variables listed in the vertical columns, enabling pairwise correlations to be presented both horizontally and vertically for the same set of variables. VPI: Video Power Index, GQS: Global quality score

across all parameters. Turkish-language YouTube videos on dental implants exhibited relatively low informational quality, with a median DISCERN score and GQS of 2. These scores suggest that most videos provided limited or insufficiently reliable content, often lacking clear explanations, balanced perspectives, or evidence-based guidance. Similar to previous studies focusing on dental implant content, the present findings indicate that Turkish-language YouTube videos often lack the necessary depth and accuracy to function as reliable educational resources.^{16,25,26} English videos consistently scored higher on both scales, suggesting better reliability and clarity. However, the absolute values were still modest, reinforcing that neither group of videos offers consistently high-quality education. This highlights a persistent gap between patient demand for online information and the actual quality of content available. This also implies that English videos not only offer more reliable information but also present it in a more patient-friendly and understandable manner. These findings are consistent with previous research indicating that English-language online health information generally tends to be of higher quality than non-English content, especially in contexts where local content production may be limited by fewer resources or lower awareness of information quality standards.^{20,21} One possible explanation for the higher quality of English-language online health information lies in the broader ecosystem of digital health communication in English-speaking countries.

Correlation analyses revealed that popularity metrics such as views, likes, and comments were closely interrelated and strongly linked with VPI in both groups. However, only English videos demonstrated consistent positive associations between quality (DISCERN, GQS) and engagement, suggesting that higher-quality content may achieve greater visibility in broader markets. In Turkish videos, the negative correlation between views and like/view ratio suggests that higher visibility did not necessarily translate into positive audience reception. These patterns reinforce the principle that popularity should not be equated with educational value. The very strong correlation observed between DISCERN and GQS scores supports the internal consistency of these quality assessment tools.

Although nearly all videos were uploaded by dentists or clinics, this professional origin did not guarantee higher quality scores, underscoring the need for structured content production standards. Specifically, English videos had a median duration nearly twice that of Turkish videos, a greater number of views, and substantially higher VPI scores. One plausible explanation is the relative scarcity of academic contributors in the Turkish digital space, coupled with lower levels of digital health literacy and fewer resources allocated to video production. These contextual factors may contribute to shorter, less reliable, and less engaging Turkish content. Interestingly, although all Turkish videos and the majority of English videos were uploaded by dentists or clinics, this professional origin did not guarantee higher quality scores, underscoring the need for structured content production standards.

Beyond methodological findings, ethical considerations deserve emphasis. Prior studies have cautioned that inaccurate or incomplete online content can mislead patients and compromise informed consent. The current results underscore the ethical responsibility of healthcare professionals and organizations to ensure that accessible, accurate, and patient-friendly materials are available. Practical strategies may include encouraging professional societies to upload peer-reviewed videos, fostering collaborations between dental professionals and content creators, and promoting patient education programs that integrate digital resources. Additionally, providing cross-language subtitles or translated versions of English videos could help mitigate information gaps for non-English-speaking populations.

Limitations

This study has several limitations that should be acknowledged. First, the cross-sectional design captures only a single time point and therefore does not account for potential temporal changes. YouTube content is dynamic, with videos being continuously uploaded, removed, or reshuffled by the platform's algorithm, which limits the reproducibility of the findings.

Second, the sample may not be fully representative of all implant-related content. Only two search terms ("dental implant" and "diş implantı") were used, in two languages (English and Turkish), and no additional keywords or alternative platforms were included. As a result, potentially relevant videos may have been overlooked. Finally, the analysis was restricted to video characteristics and quality scoring. Patient comprehension, information retention, and real-world behavioral outcomes were not evaluated, which limits conclusions about the true educational impact of the videos.

CONCLUSION

As a result, given the widespread use of YouTube by patients seeking information about dental procedures such as implant placement, the discrepancy in information quality between English and Turkish content is concerning. However, the broader message is that both groups demonstrated only suboptimal educational quality, consistent with previous research showing that most medical and dental videos on YouTube fall within poor-to-moderate reliability ranges. Although English-language videos performed moderately better than Turkish ones, their absolute DISCERN and GQS scores indicate that they cannot be considered consistently reliable sources of patient education. This highlights the urgent need for dental professionals and organizations, both locally and internationally, to create high-quality, evidence-based video content. Using standardized templates and collaborating with structured patient education programs could enhance both accuracy and accessibility. Moreover, strategies to bridge language gaps-such as adding subtitles to English videos or developing cross-language versions-should be prioritized to reduce disparities in information access. Finally, expanding digital health literacy initiatives will empower patients to critically evaluate the quality of online

information, ultimately supporting more informed decision-making.

ETHICAL DECLARATIONS

Ethics Committee Approval

Since this study did not involve human participants or data derived from patients and was conducted using publicly accessible online content, approval from an ethics committee was not required.

Informed Consent

As no human subjects were involved in the study and the data were collected from publicly available YouTube videos, informed consent was not applicable.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Monje A, Alcoforado G, Padial-Molina M, Suarez F, Lin GH, Wang HL. Generalized aggressive periodontitis as a risk factor for dental implant failure: a systematic review and meta-analysis. *J Periodontol*. 2014; 85(10):1398-1407. doi:10.1902/jop.2014.140135
- Giannobile WV, Lang NP. Are dental implants a panacea or should we better strive to save teeth? *J Dent Res*. 2016;95(1):5-6. doi:10.1177/0022034515618942
- Chrcanovic BR, Kisch J, Albrektsson T, Wennerberg A. Survival of dental implants placed in sites of previously failed implants. *Clin Oral Implants Res*. 2017;28(11):1348-1353. doi:10.1111/clr.12992
- Abukaraky A, Hamdan AA, Ameera MN, Nasief M, Hassona Y. Quality of YouTube TM videos on dental implants. *Med Oral Patol Oral Cir Bucal*. 2018;23(4):e463-e468. doi:10.4317/medoral.22447
- Wang G, Gao X, Lo EC. Public perceptions of dental implants: a qualitative study. *J Dent*. 2015;43(7):798-805. doi:10.1016/j.jdent.2015.04.012
- Nason K, Donnelly A, Duncan HF. YouTube as a patient-information source for root canal treatment. *Int Endod J*. 2016;49(12):1194-1200. doi:10.1111/iej.12575
- Madathil KC, Rivera-Rodriguez AJ, Greenstein JS, Gramopadhye AK. Healthcare information on YouTube: a systematic review. *Health Informatics J*. 2015;21(3):173-194. doi:10.1177/1460458213512220
- Alzoubi H, Karasneh R, Irshaidat S, et al. Exploring the use of YouTube as a pathology learning tool and its relationship with pathology scores among medical students: cross-sectional study. *JMIR Med Educ*. 2023;9:e45372. doi:10.2196/45372
- Díaz-Rodríguez A, Limeres-Posse J, Albuquerque R, et al. Assessment of the quality of oral biopsy procedure videos shared on YouTube. *Oral Dis*. 2024;30(5):3081-3093. doi:10.1111/odi.14690
- Gross RT, Ghaltakhchyan N, Nanney EM, et al. Evaluating video-based lectures on YouTube for dental education. *Orthod Craniofac Res*. 2023; 26(Suppl 1):210-220. doi:10.1111/ocr.12669
- Romano A, Fiori F, Petruzzi M, Della Vella F, Serpico R. YoutubeTM content analysis as a means of information in oral medicine: a systematic review of the literature. *Int J Environ Res Public Health*. 2022;19(9):5451. doi:10.3390/ijerph19095451
- Ventura JVL, da Silva CF, Lopes TDS, et al. Ameloblastoma on YouTube: content analysis and strategies to address misinformation. *Oral Maxillofac Surg*. 2025;29(1):91. doi:10.1007/s10006-025-01381-z
- Bulut SO, Ercel NO. Quality of free gingival graft content in youtube videos: Usability in patient information and student education. *Med Oral Patol Oral Cir Bucal*. 2023;28(6):e607-e613. doi:10.4317/medoral.26011
- Hassona Y, Taimeh D, Marahleh A, Scully C. YouTube as a source of information on mouth (oral) cancer. *Oral Dis*. 2016;22(3):202-208. doi:10.1111/odi.12434
- Menziletoglu D, Guler AY, Isik BK. Are YouTube videos related to dental implant useful for patient education? *J Stomatol Oral Maxillofac Surg*. 2020;121(6):661-664. doi:10.1016/j.jormas.2019.12.022
- Tayan E, Gökmenoglu C, Büyük SK. YouTube™ as an education platform for dental implants. *J Int Dent Sci*. 2024;10(3):159-166. doi:10.21306/dishkekimligi.1456647
- Di Spirito F, Giordano F, Di Palo MP, et al. Reliability and accuracy of YouTube peri-implantitis videos as an educational source for patients in population-based prevention strategies. *Healthcare (Basel)*. 2023;11(14):2094. doi:10.3390/healthcare11142094
- Ince Kuka G, Gursoy H. Content quality and reliability of YouTube™ videos as a source of information about good oral hygiene practices in adults. *PeerJ*. 2024;12:e18183. doi:10.7717/peerj.18183
- 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(2):105-111. doi:10.1136/jech.53.2.105
- Yelbay M, Karacaoglu F, Kurumlu M. Evaluation of the quality and reliability of Turkish and English YouTube video contents on gingival recession treatment. *Sci Rep*. 2025;15(1):14444. doi:10.1038/s41598-025-99371-4
- Berland GK, Elliott MN, Morales LS, et al. Health information on the Internet: accessibility, quality, and readability in English and Spanish. *JAMA*. 2001;285(20):2612-2621. doi:10.1001/jama.285.20.2612
- Başkan HK, Başkan B. YouTube as a source of information on dental trauma: a quality, accuracy, and reliability analysis. *Dent Traumatol*. 2025;0:1-14. doi:10.1111/edt.13088
- Radonjic A, Fat Hing NN, Harlock J, Naji F. YouTube as a source of patient information for abdominal aortic aneurysms. *J Vasc Surg*. 2020; 71(2):637-644. doi:10.1016/j.jvs.2019.08.230
- Sampson M, Cumber J, Li C, Pound CM, Fuller A, Harrison D. A systematic review of methods for studying consumer health YouTube videos, with implications for systematic reviews. *PeerJ*. 2013;1:e147. doi:10.7717/peerj.147
- Duygu G. YouTube videolarının dental implant içeriği hastalara faydalı mı? *7tepe Klin Derg*. 2024;20(1):35-42. doi:10.5505/yeditepe.2024.99266
- Bicer AY, Zor ZF, Bağcı N, Peker İ. İmplant konulu YouTube videolarının değerlendirilmesi. *ADO Klin Bil Derg*. 2021;10(3):186-192.

Immunonutritional profiling for prognostic assessment in diffuse large b-cell lymphoma: an initial evaluation of the C-reactive Protein-Albumin-Lymphocyte (CALLY) Index

İbrahim Ethem Pınar¹, Vildan Özkocaman¹, Tuba Ersal¹, Vildan Gürsoy¹,
Fazıl Çağrı Hunutlu¹, Elif Yiğit Ayhan², Fahir Özkalemkaş¹

¹Department of Hematology, Department of Internal Medicine, Faculty of Medicine, Bursa Uludağ University, Bursa, Türkiye

²Department of Internal Medicine, Faculty of Medicine, Bursa Uludağ University, Bursa, Türkiye

Cite this article as: Pınar İE, Özkocaman V, Ersal T, et al. Immunonutritional profiling for prognostic assessment in diffuse large b-cell lymphoma: an initial evaluation of the C-reactive Protein-Albumin-Lymphocyte (CALLY) Index. *J Health Sci Med.* 2025;8(6):1112-1117.

Received: 23.08.2025

Accepted: 10.10.2025

Published: 25.10.2025

ABSTRACT

Aims: Diffuse large B-cell lymphoma (DLBCL) demonstrates wide heterogeneity, complicating risk stratification with conventional prognostic instruments, particularly the International Prognostic Index (IPI). The C-reactive Protein (CRP)-Albumin-Lymphocyte (CALLY) Index, an immunonutritional marker that integrates inflammation, nutritional status, and immune competence, has demonstrated prognostic relevance in solid tumors. The present analysis was designed to determine the clinical utility of the CALLY Index for prognostication in newly diagnosed DLBCL patients and to assess its added value when integrated with traditional prognostic scores.

Methods: In this retrospective cohort, 112 patients presenting with newly diagnosed DLBCL and treated at Bursa Uludağ University between 2015 and 2019 were evaluated. The CALLY Index was calculated as (serum albumin × absolute lymphocyte count) / (CRP × 10⁴). An optimal cutoff value of 0.78 was derived using log-rank testing and used to stratify patients into low (≤0.78) and high (>0.78) CALLY categories. Baseline clinical features were compared between groups. Kaplan-Meier curves and multivariable Cox proportional hazards modeling were applied to examine overall survival (OS). Multivariable analysis included IPI score and serum beta-2-microglobulin.

Results: Cases in the low CALLY subgroup (68.8%) were significantly older, had poorer Eastern Cooperative Oncology Group performance, lower albumin, higher CRP, and more advanced disease (p<0.01 for all). They were also more frequently classified into higher-risk IPI categories (p=0.009). Median OS in the low CALLY group was 14.8 months, while the high CALLY group had not reached median OS at the end of follow-up (p=0.0009). Univariable analysis revealed that low CALLY (HR: 5.33, p=0.002) and high IPI score were associated with worse OS. In multivariable analysis, low CALLY remained an independent predictor of mortality (HR: 3.42, 95% CI: 1.16-10.08; p=0.025), even after adjusting for IPI and beta-2-microglobulin levels.

Conclusion: The CALLY Index is an independent and clinically accessible prognostic biomarker in DLBCL. Its integration of inflammatory, nutritional, and immune parameters provides complementary prognostic information beyond traditional models such as the IPI. Given its cost-effectiveness and reliance on routine laboratory data, the CALLY Index may serve as a valuable tool in real-world prognostic assessment. These findings support prospective validation and exploration of its utility in dynamic risk models and personalized treatment strategies for DLBCL.

Keywords: Diffuse large B-cell lymphoma, CALLY Index, immunonutritional profiling, prognosis, International Prognostic Index, overall survival

INTRODUCTION

Diffuse large B-cell lymphoma (DLBCL), the predominant form of non-Hodgkin lymphoma, is marked by substantial heterogeneity at the clinical, pathological, and molecular levels. Although standard chemoimmunotherapy-most notably the R-CHOP regimen (rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisone)-has significantly improved treatment outcomes, survival rates remain variable. This variability underscores the pressing need for more refined and biologically informed prognostic models.

The International Prognostic Index (IPI) continues to serve as the predominant clinical instrument for risk stratification in DLBCL. However, its performance in the rituximab era has been questioned due to its limited ability to discriminate high-risk patients and the large proportion of individuals falling into intermediate-risk groups, which restricts its clinical utility.¹⁻³ Moreover, IPI does not incorporate molecular or biomarker-based parameters, which have shown promise in enhancing risk prediction.⁴ Alternative models such as

the National Comprehensive Cancer Network (NCCN)-IPI, revised (R)-IPI, and biomarker-integrated indices (e.g., double expressor status, MYC/BCL2 coexpression) offer improved stratification but are often complex or not readily available in routine clinical practice.⁵⁻⁷

Recent investigations highlight the prognostic relevance of systemic inflammation, nutritional state, and immune competence in DLBCL.⁸⁻¹⁰ Our group has also investigated systemic inflammatory and immunonutritional indices in hematologic malignancies, including myelofibrosis, Hodgkin lymphoma, myelodysplastic syndromes, and primary central nervous system (CNS) lymphoma, highlighting their potential in prognostic assessment.¹¹⁻¹⁴ Inflammatory and immunonutritional indices such as the Prognostic Nutritional Index (PNI), glasgow prognostic score (GPS), neutrophil-to-lymphocyte ratio (NLR), and platelet-to-lymphocyte ratio (PLR) have demonstrated independent associations with survival outcomes and have been proposed as supplements to classical models.^{8,15-17} However, each of these indices captures only a single biological dimension.

The C-reactive Protein (CRP)-Albumin-Lymphocyte (CALLY) Index is a novel composite biomarker that integrates inflammation (CRP), nutritional status (serum albumin), and immunity (absolute lymphocyte count). The CALLY Index has shown strong independent prognostic value across various solid tumors, including gastric, colorectal, esophageal, hepatocellular, lung, breast, and ovarian cancers.¹⁸⁻²² Compared to single-parameter models, it offers a more comprehensive biological assessment and has outperformed traditional markers such as CRP or albumin alone in several malignancies.^{18,20,22}

The appeal of the CALLY Index lies in its simplicity, cost-effectiveness, and reliance on readily accessible laboratory parameters, making it feasible for use in daily clinical practice and in resource-limited settings. While its usefulness has been documented in solid cancers, the prognostic role of the CALLY Index remains unexplored in hematologic neoplasms, including DLBCL.¹⁸

Our primary aim was to determine the prognostic relevance of the CALLY Index in the setting of newly diagnosed DLBCL. We hypothesize that this composite biomarker, encompassing inflammatory, nutritional, and immune parameters, may serve as a reliable and accessible adjunct to current clinical models, ultimately supporting more personalized risk stratification strategies.

METHODS

Ethical Considerations

This study was conducted in accordance with the principles of the Declaration of Helsinki and was approved by the Bursa Uludağ University Faculty of Medicine Clinical Researches Ethics Committee (Decision No: 2021-7/30, Date: 02.06.2021).

Study Design and Patient Selection

The study retrospectively analyzed individuals diagnosed with DLBCL at Bursa Uludağ University, Division of Hematology, between January 2015 and December 2019. Eligibility

criteria were age ≥ 18 years, availability of baseline laboratory parameters for CALLY Index calculation, and accessible survival data. Patients with incomplete laboratory or clinical information were excluded. Altogether, 112 individuals satisfying these requirements were analyzed in the study.

Data Collection and Definitions

Baseline clinical and laboratory information at diagnosis was retrieved from the institutional electronic database. Variables collected included patient age, sex, serum albumin (g/dl), absolute lymphocyte count (cells/ μ L), CRP (mg/dl), Eastern Cooperative Oncology Group performance status (ECOG PS), serum lactate dehydrogenase (LDH), beta-2-microglobulin concentration, number of extranodal sites involved, Ann Arbor stage, and IPI score. The CALLY Index was calculated as:

$$\frac{\text{Serum albumin (g/dl)} \times \text{absolute lymphocyte count (/}\mu\text{L)}}{\text{CRP (mg/dl)} \times 10^4}$$

CALLY Stratification

The optimal stratification point of the CALLY Index for predicting overall survival (OS) was determined using log-rank test p-values across a range of potential thresholds. Based on this analysis, a CALLY score of 0.78 was selected as the discriminative cutoff (Figure 1). Patients were subsequently categorized into low (≤ 0.78) and high (> 0.78) CALLY groups.

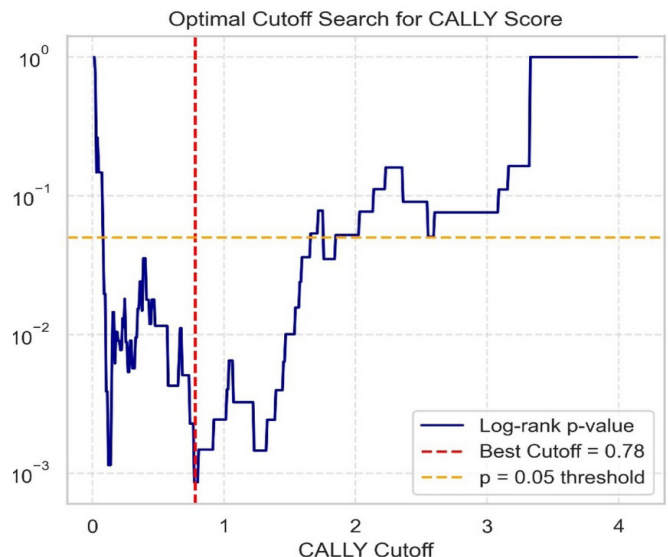


Figure 1. Optimal CALLY Index cutoff determination using log-rank test p-values

Blue Line: Log-rank p-values calculated at each potential CALLY cutoff. Red Dashed Line: Best discriminative cutoff for overall survival (CALLY=0.78). Orange Dashed Line: Statistical significance threshold ($p=0.05$)

CALLY: C-reactive Protein-Albumin-Lymphocyte Index

Statistical Analysis

Continuous variables were described using medians with corresponding interquartile ranges (IQR) and their distributions were compared through the Mann-Whitney U test. Categorical variables were presented as counts and percentages, with group differences assessed using either the Chi-square test or Fisher's exact test, depending on suitability.

OS was defined as the interval between the date of diagnosis and either death from any cause or the most recent follow-up. Survival distributions were estimated by the Kaplan–Meier method, and differences between groups were examined using the log-rank test. Both univariate and multivariable Cox proportional hazards regression models were applied to explore the impact of clinical variables on OS. Candidate variables with a p-value <0.10 in univariate analysis were subsequently incorporated into the multivariable model. Effect sizes were expressed as hazard ratios (HRs) together with 95% confidence intervals (CIs).

All statistical procedures were performed in Python (version 3.13.2) using the lifelines package. A two-sided p-value <0.05 was considered to indicate statistical significance.

RESULTS

The final analysis comprised 112 patients with DLBCL for whom CALLY index values could be calculated. At the time of diagnosis, the study cohort had a median age of 55 years (IQR: 43.8–66.0), with an overall range of 18 to 89 years. Males represented the majority of the study population (60.6%, n=67).

Baseline Laboratory and Clinical Characteristics

At baseline, patients exhibited a median serum albumin level of 3.8 g/dl (IQR: 3.4–4.3), median absolute lymphocyte count of 1575/μL (IQR: 953–2238), and median CRP level of 2.2 mg/L (IQR: 0.5–7.2). The resulting median CALLY Index was 0.2 (IQR: 0.1–1.3), reflecting a broad spectrum of inflammatory and nutritional status across the cohort (Figure 2).

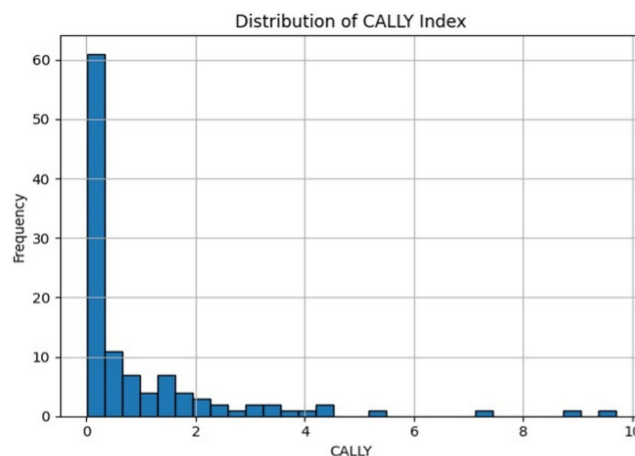


Figure 2. Histogram of CALLY Index distribution among DLBCL patients

The CALLY Index is calculated as: (serum albumin × lymphocyte count) / (CRP × 10⁴). Values are right-skewed, with most patients exhibiting low scores

CALLY: C-reactive Protein–Albumin–Lymphocyte Index, DLBCL: Diffuse large B-cell lymphoma, CRP: C-reactive protein

Patients were stratified into low CALLY (≤0.78, n=77, 68.8%) and high CALLY (>0.78, n=35, 31.3%) groups based on the optimal cutoff derived from survival analyses.

Comparison Between CALLY Groups

Individuals in the low CALLY cohort were significantly older (median age: 59.0 vs. 48.0 years, p=0.005), had lower serum albumin (3.7 vs. 4.4 g/dl, p<0.001), lower lymphocyte counts (1360/μL vs. 1980/μL, p<0.001), and markedly elevated CRP levels (4.6 vs. 0.3 mg/dl, p<0.001) compared to those in the high CALLY group (Table 1).

Table 1. Baseline clinical and laboratory characteristics of patients with DLBCL stratified by CALLY Index (cut-off: 0.78)

Variable	All patients (n=112)	Low CALLY (≤0.78) n=77	High CALLY (>0.78) n=35	p-value
Age, years, median (IQR)	55.0 (43.8–66.0)	59.0 (47.0–70)	48.0 (38.5–56.5)	0.005 ^a
Sex, male, n (%)	65 (58.0%)	46 (59.7%)	19 (54.3%)	0.737 ^b
Albumin, g/dl, median (IQR)	3.8 (3.4–4.3)	3.7 (3.1–4.0)	4.4 (4.1–4.5)	<0.001 ^a
Lymphocyte count (/μL), median (IQR)	1575 (953–2238)	1360 (730–1950)	1980 (1540–2525)	<0.001 ^a
CRP, mg/dl, median (IQR)	2.2 (0.5–7.2)	4.6 (2.1–11.0)	0.3 (0.3–0.6)	<0.001 ^a
Serum beta-2 microglobulin, normalized ratio, n (%)				0.003 ^b
≤1		28 (36.4%)	24 (68.6%)	
>1		49 (63.6%)	11 (31.4%)	
ECOG PS, median (IQR)	2 (1–3)	2 (1–3)	1 (1–2)	0.026 ^a
Extranodal site count, median (IQR)	1 (0–2)	1 (1–2)	1 (0–1)	0.001 ^a
Serum LDH, normalized ratio, n (%)				0.024 ^b
≤1		25 (32.5%)	20 (57.1%)	
>1		52 (67.5%)	15 (42.9%)	
Ann arbor stage, median (IQR)	3 (2–4)	4 (2–4)	3 (2–4)	0.008 ^a
IPI group distribution, n (%)				0.009 ^b
Low		15 (23.4%)	16 (50.0%)	
Low-intermediate		9 (14.1%)	7 (21.9%)	
High-intermediate		17 (26.6%)	6 (18.8%)	
High		23 (35.9%)	3 (9.4%)	
CALLY Index, median (IQR)	0.2 (0.1–1.3)	0.12 (0.04–0.27)	2.03 (1.46–3.44)	<0.001 ^a

DLBCL: Diffuse large B-cell lymphoma, CALLY: C-reactive Protein–Albumin–Lymphocyte, IQR: Interquartile range, CRP: C-reactive protein, ECOG PS: Eastern Cooperative Oncology Group Performance Status, LDH: Lactate dehydrogenase, IPI: International Prognostic Index, a: Mann–Whitney U test, b: Chi-square or Fisher's exact test

Furthermore, the median ECOG PS was poorer in the low CALLY category [2 (IQR: 1-3)] than in the high CALLY category [1 (IQR: 1-2), $p=0.026$]. Similarly, the extent of extranodal involvement was greater [median: 1 (IQR: 1-2) vs. 1 (IQR: 0-1), $p=0.001$] (Table 1).

Analysis of Ann Arbor staging revealed a higher stage distribution in the low CALLY group [median: 4 (IQR: 2-4)] versus the high CALLY group [3 (IQR: 2-4), $p=0.008$] (Table 1).

Risk Stratification by IPI Score

Distribution of patients according to the IPI revealed a significantly greater representation of higher-risk categories in the low CALLY group ($p=0.009$) (Table 1 and Figure 3). Risk stratification revealed that 71.9% of individuals in the high CALLY group were categorized as low or low-intermediate risk, compared with 37.5% among those with low CALLY scores. Conversely, 62.5% of patients in the low CALLY group fell into the high-intermediate or high-risk categories, whereas this proportion was only 28.1% among patients in the high CALLY group (Figure 3).

Collectively, these findings suggest that a lower CALLY Index at diagnosis is significantly associated with older age, poor nutritional and inflammatory profiles, greater tumor burden, and more unfavorable prognostic features, as reflected by ECOG PS, extranodal site involvement, Ann Arbor stage, and IPI risk group distribution.

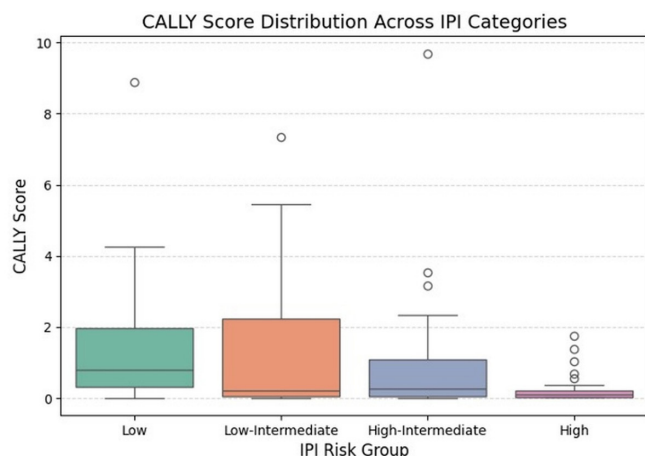


Figure 3. Distribution of CALLY scores across IPI risk groups

IPI categories: Low, low-intermediate, high-intermediate, high. Boxes: IQRs. Lines inside boxes: Medians. Whiskers: $1.5 \times \text{IQR}$. Circles: Outliers

CALLY: C-reactive Protein-Albumin-Lymphocyte Index, IPI: International Prognostic Index, IQR: Interquartile range

Survival Analysis by CALLY Score

Patients with CALLY values ≤ 0.78 demonstrated significantly worse OS than those with scores above 0.78, as illustrated by Kaplan-Meier survival analysis. Specifically, the median OS for the low CALLY group was 14.8 months, whereas it was not reached for the high CALLY group within the follow-up period, suggesting superior long-term survival in this subgroup. The survival curves began to diverge early during follow-up and remained consistently separated over time, indicating a sustained prognostic distinction. OS varied significantly between groups, with the log-rank test confirming statistical significance ($p=0.0009$) (Figure 4). The results provide evidence for the prognostic significance of the CALLY Index at diagnosis in stratifying survival outcomes among patients with DLBCL.

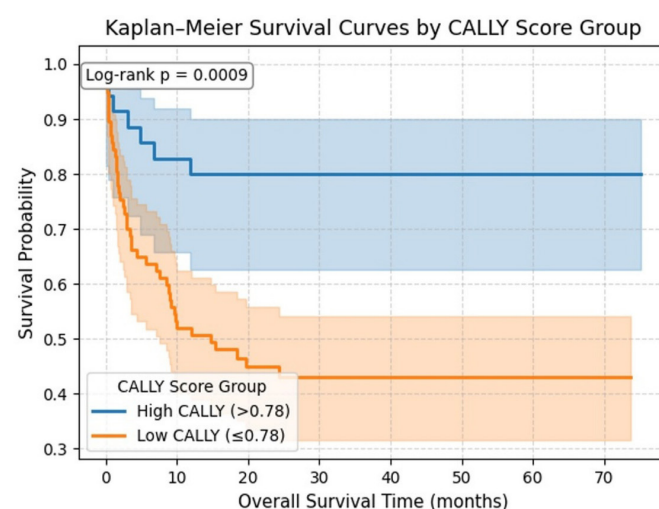


Figure 4. Kaplan-Meier overall survival curves according to CALLY Index stratification

Blue line: High CALLY (>0.78). Orange line: Low CALLY (≤ 0.78). Shaded areas: 95% confidence intervals. Log-rank $p=0.0009$

CALLY: C-reactive Protein-Albumin-Lymphocyte Index

Cox Proportional Hazards Analysis

In univariable Cox regression analysis, both IPI risk category and low CALLY Index were significant predictors of OS. Specifically, cases in the low CALLY category had significantly worse survival compared to those with high CALLY (HR: 5.33, 95% CI: 1.88-15.12; $p=0.002$). Similarly, increasing IPI category showed a significant relationship with increased mortality, with the high-risk group showing a particularly elevated hazard (HR: 14.02, 95% CI: 4.08-48.26; $p<0.001$) (Table 2).

Table 2. Cox regression analysis of factors associated with overall survival in DLBCL patients

Variable	Univariate HR (95% CI)	Univariate p	Multivariate HR (95% CI)	Multivariate p
CALLY score (low)	5.33 (1.88-15.12)	0.002	3.42 (1.16-10.08)	0.025
IPI (low-intermediate)	3.25 (0.73-14.52)	0.123	0.27 (0.09-0.85)	0.025
IPI (high-intermediate)	7.69 (2.14-27.59)	0.002	0.54 (0.25-1.17)	0.118
IPI (high)	14.02 (4.08-48.26)	<0.001	0.07 (0.02-0.26)	<0.001
Beta-2-microglobulin (high)	1.85 (0.94-3.61)	0.073	0.59 (0.28-1.23)	0.156

HRs and 95% CIs were calculated using Cox proportional hazards regression. Univariable and multivariable analyses were performed to assess the prognostic significance of the CALLY score, IPI categories, and beta-2 microglobulin levels. In the multivariable model, both low CALLY score and higher IPI category remained independent predictors of poor overall survival. A low CALLY score was defined as ≤ 0.78 . IPI categories were referenced against the low-risk group.

DLBCL: Diffuse large B-cell lymphoma, HR: Hazard ratio, CI: Confidence interval, CALLY: C-reactive Protein-Albumin-Lymphocyte Index, IPI: International Prognostic Index

In the multivariable Cox model adjusting for IPI risk categories and beta-2-microglobulin, low CALLY remained independently associated with increased mortality (HR: 3.42, 95% CI: 1.16-10.08; $p=0.025$) (Table 2). Notably, the inclusion of the CALLY Index in the multivariable model did not attenuate the prognostic impact of the IPI, indicating that both metrics offer complementary prognostic information.

The sustained statistical significance of the CALLY Index in multivariable Cox analysis, even after adjusting for IPI risk categories, underscores its ability to capture prognostic dimensions not fully addressed by the IPI. The study highlights the potential of the CALLY Index as a clinically applicable, cost-efficient, and non-invasive biomarker for improving risk categorization and tailoring prognostic assessment in DLBCL patients.

DISCUSSION

This work constitutes the initial attempt to explore the prognostic implications of the CALLY Index in DLBCL. The present analysis indicates that a low baseline CALLY Index independently correlates with inferior OS, even when adjusted for established clinical prognostic tools such as the IPI and the NCCN-IPI. These results are consistent with previously reported associations between systemic inflammatory/nutritional biomarkers and outcomes in both solid tumors and lymphomas.^{8,23}

Conventional prognostic models such as the IPI and NCCN-IPI exhibit limited capacity to capture the biological heterogeneity of DLBCL, particularly among patients classified within intermediate-risk categories.²⁴ This limitation has driven the development of complementary or alternative strategies incorporating immune-nutritional biomarkers-such as the PNI, Systemic Immune-Inflammation Index (SII), and Advanced Lung Cancer Inflammation Index-each of which contributes to a more nuanced risk stratification framework. In recent years, growing attention has increasingly focused on the intricate interplay of systemic inflammation, nutritional state, and tumor progression as a critical determinant of cancer prognosis. Within the context of DLBCL, biomarkers such as the NLR, SII, PLR, and PNI have been independently associated with prognostic outcomes.^{10,25-27} Among these, the CALLY Index-comprising serum albumin, absolute lymphocyte count, and CRP-offers a unified reflection of immunonutritional status and systemic inflammation, thus capturing a more integrative measure of tumor-host interactions.²⁸

Moreover, dynamic biomarkers such as circulating tumor deoxyribonucleic acid (ctDNA) provide real-time insights into disease burden and therapeutic response. Studies have shown that early ctDNA clearance after treatment initiation predicts better survival, even before radiologic response is evident.^{29,30} However, ctDNA measurement is limited by cost and availability, making readily accessible laboratory indices such as CALLY a more practical option for most institutions. In line with this, our prior investigations in DLBCL and primary CNS lymphoma have underscored the clinical utility of easily obtainable laboratory markers for real-world prognostic assessment.^{14,31,32}

In our study, the CALLY Index maintained its prognostic significance across clinically relevant subgroups, including patients stratified by age, stage, extranodal involvement, and LDH levels. Its reliance on routine laboratory values makes it highly feasible for use in real-world settings. Importantly, the CALLY Index was independently predictive of OS in multivariate Cox models, reinforcing its potential as a robust prognostic biomarker.

Limitations

Among the study's limitations, its retrospective and monocentric design stands out as a major factor limiting external generalizability. Moreover, dynamic alterations in the CALLY Index and their association with therapeutic response were not investigated. External validation across independent patient cohorts remains essential to confirm the prognostic robustness of the CALLY Index in DLBCL.

Future investigations should prioritize prospective, multicenter cohorts to confirm the prognostic significance of the CALLY Index. Additionally, integrating CALLY with imaging-based response assessments, molecular biomarkers, and measurable residual disease markers may enhance precision prognostication. Investigating how changes in the CALLY Index during therapy correlate with therapeutic response or relapse could also provide clinically actionable insights. Finally, studies exploring the prognostic impact of the CALLY Index in novel therapeutic settings, including immunotherapy or bispecific antibody-based regimens, would further expand its applicability.

CONCLUSION

In summary, the CALLY Index serves as an independent and readily applicable prognostic indicator of OS in DLBCL. By incorporating markers of systemic inflammation and nutritional status, it provides added value to traditional clinical models such as IPI. Its simplicity, cost-effectiveness, and widespread availability render it a practical tool in routine practice. Well-designed prospective trials are warranted to substantiate its prognostic utility and to clarify its role in shaping personalized management strategies.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Bursa Uludağ University Faculty of Medicine Clinical Researches Ethics Committee (Decision No: 2021-7/30, Date: 02.06.2021).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Bicler J, Eloranta S, de Nully Brown P, et al. Simplicity at the cost of predictive accuracy in diffuse large B-cell lymphoma: a critical assessment of the R-IPI, IPI, and NCCN-IPI. *Cancer Medicine*. 2018;7(1):114-122. doi:10.1002/cam4.1271
- Sehn LH, Berry B, Chhanabhai M, et al. The revised International Prognostic Index (R-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Blood*. 2006;109(5):1857-1861. doi:10.1182/blood-2006-08-038257
- Spiegel JY, Cheung MC, Guirguis HR, Buckstein R. Validation of the NCCN-IPI in both de novo and transformed diffuse large B cell lymphoma. *Leukemia Lymphoma*. 2017;58(1):214-217. doi:10.1080/10428194.2016.1179295
- Mu S, Shi D, Ai L, et al. International Prognostic Index-based immune prognostic model for diffuse large B-cell lymphoma. *Front Immunol*. 2021;12:732006. doi:10.3389/fimmu.2021.732006
- Wang J, Zhou M, Xu JY, et al. MYC and BCL-2 adjusted-International Prognostic Index (A-IPI) is a better predictor of outcome than the standard IPI for patients with diffuse large B-cell lymphoma treated with R-CHOP. *Histol Histopathol*. 2016;31(3):285-292. doi:10.14670/hh-11-673
- Khanmohammadi S, Masrouf M, Fallahtafi P, Hasani F. MicroRNA as a potential diagnostic and prognostic biomarker in diffuse large B-cell lymphoma: a systematic review and meta-analysis. *Cancer Reports*. 2025;8(1):e70070. doi:10.1002/cnr2.70070
- Perry AM, Mitrovic Z, Chan WC. Biological prognostic markers in diffuse large B-cell lymphoma. *Cancer Control*. 2012;19(3):214-226. doi:10.1177/107327481201900306
- Peng Y, Jiang T, Chen S, et al. Nutritional and immune-inflammatory scoring system for predicting prognosis in patients with newly diagnosed diffuse large B-cell lymphoma. *Blood*. 2024;144(Supplement 1):6463-6463. doi:10.1182/blood-2024-204285
- He J, Yin H, Xia Y, et al. Prognostic Nutritional Index, a novel biomarker which predicts worse prognosis in diffuse large B cell lymphoma. *Leukemia Research*. 2021;110:106664. doi:10.1016/j.leukres.2021.106664
- Waley AB, Haggag R, Ahmed Barakat AA, et al. Impact of Prognostic Nutritional Index and Systemic Immune-Inflammation Index on the clinical outcome of diffuse large B cell lymphoma patients treated with RCHOP. *Zagazig Univ Med J*. 2024;30(8.1):3907-3917. doi:10.21608/zumj.2024.256794.3059
- Ersal T, Özkocaman V, Pinar İE, et al. Systemic inflammatory indices for predicting prognosis of myelofibrosis. *Sci Rep*. 2023;13(1):12539. doi:10.1038/s41598-023-39077-7
- Gürsoy V, Hunutlu F, Pinar İE, et al. The clinical impacts of the controlling nutritional status score on patients with Hodgkin lymphoma. *Eur Rev Med Pharmacol Sci*. 2023;27(20):9916-9927. doi:10.26355/eurrev_202310_34170
- Gursoy V, Sadri S, Kucukelyas HD, et al. HALP score as a novel prognostic factor for patients with myelodysplastic syndromes. *Sci Rep*. 2024;14(1):13843. doi:10.1038/s41598-024-64166-6
- Ersal T, Özkocaman V, Pinar İE, et al. Potential prognostic parameters and real-world data in patients with primary central nervous system lymphoma: a new brick on the old ones. *Eur Res J*. 2023;9(5):1157-1165. doi:10.18621/eurj.1267903
- Hao X, Wei Y, Wei X, et al. Glasgow prognostic score is superior to other inflammation-based scores in predicting survival of diffuse large B-cell lymphoma. *Oncotarget*. 2017;8(44):76740-76748. doi:10.18632/oncotarget.20832
- Liu T, Ye F, Li Y, Liu A. Comparison and exploration of the prognostic value of the Advanced Lung Cancer Inflammation Index, Prognostic Nutritional Index, and Systemic Immune-Inflammation Index in newly diagnosed diffuse large B-cell lymphoma. *Ann Palliat Med*. 2021;10(9):9650-9659.
- Go S-I, Park S, Kang MH, Kim H-G, Kim HR, Lee G-W. Clinical impact of Prognostic Nutritional Index in diffuse large B cell lymphoma. *Ann Hematol*. 2019;98(2):401-411. doi:10.1007/s00277-018-3540-1
- Chen D, Ma Y, Li J, et al. Prognostic and clinicopathological significance of C-reactive protein-albumin-lymphocyte(CALLY) in patients with digestive system neoplasms: a systematic review and meta-analysis. *World J Surg Oncol*. 2025;23(1):114. doi:10.1186/s12957-025-03779-1
- Liu X-Y, Zhang X, Zhang Q, et al. The value of CRP-Albumin-Lymphocyte Index (CALLY Index) as a prognostic biomarker in patients with non-small cell lung cancer. *Support Care Cancer*. 2023;31(9):533. doi:10.1007/s00520-023-07997-9
- Feng J, Wang L, Yang X, Chen Q. Clinical significance of preoperative CALLY Index for prognostication in patients with esophageal squamous cell carcinoma undergoing surgery. *Sci Rep*. 2024;14(1):713. doi:10.1038/s41598-023-51109-w
- Fukushima N, Masuda T, Tsuboi K, et al. Prognostic significance of the preoperative C-reactive protein-albumin-lymphocyte (CALLY) index on outcomes after gastrectomy for gastric cancer. *Surgery Today*. 2024;54(8):943-952. doi:10.1007/s00595-024-02813-1
- Aoyama T, Maezawa Y, Hashimoto I, et al. The CRP-Albumin-Lymphocyte (CALLY) Index is an independent prognostic factor for gastric cancer patients who receive curative treatment. *Anticancer Res*. 2024;44(4):1629. doi:10.21873/anticancer.16961
- Wang Z, Zhang J, Luo S, Zhao X. Prognostic significance of Systemic Immune-Inflammation Index in patients with diffuse large B-cell lymphoma. *Front Oncol*. 2021;11: 655259. doi:10.3389/fonc.2021.655259
- Jelicic J, Juul-Jensen K, Bukumiric Z, et al. Prognostic indices in diffuse large B-cell lymphoma: a population-based comparison and validation study of multiple models. *Blood Cancer J*. 2023;13(1):157. doi:10.1038/s41408-023-00930-7
- Wu J, Zhu H, Zhang Q, et al. Nomogram based on the systemic immune-inflammation index for predicting the prognosis of diffuse large B-cell lymphoma. *Asia-Pacific J Clin Oncol*. 2023;19(2):e138-e148. doi:10.1111/ajco.13806
- Shen Z, Wang F, He C, et al. The Value of Prognostic Nutritional Index (PNI) on newly diagnosed diffuse large B-cell lymphoma patients: a multicenter retrospective study of HHLWG based on propensity score matched analysis. *J Inflamm Res*. 2021;14:5513-5522. doi:10.2147/jir.s340822
- Su F, Lian K. Prognostic evaluation of System Immune-Inflammatory Index and Prognostic Nutritional Index in double expressor diffuse large B-cell lymphoma. *Open Medicine*. 2023;18(1):20230819. doi:10.1515/med-2023-0819
- Iida H, Tani M, Komeda K, et al. Superiority of CRP-Albumin-Lymphocyte Index (CALLY Index) as a non-invasive prognostic biomarker after hepatectomy for hepatocellular carcinoma. *HPB*. 2022;24(1):101-115. doi:10.1016/j.hpb.2021.06.414
- Cherng H-JJ, Herrera A. Circulating Tumor DNA in Diffuse Large B-Cell Lymphoma: from Bench to Bedside? *Curr Treat Option Oncol*. 2024;25(5):659-678. doi:10.1007/s11864-024-01201-8
- Fu L, Zhou X, Zhang X, et al. Circulating tumor DNA in lymphoma: technologies and applications. *J Hematol Oncol*. 2025;18(1):29. doi:10.1186/s13045-025-01673-7
- Pinar İE, Özkocaman V, Ersal T, et al. Comparison of international prognostic indices and validation study for patients with diffuse large B-cell lymphoma in the Rituximab Era. Article. *Uluslararası Hematol Onkol Derg*. 2023;33(2):57-65. doi:10.4999/uhod.236953
- Demirci U, Kırkızlar HO, Ümit EG, et al. CD5 as a prognostic marker in patients with diffuse large B-cell lymphoma: a multicenter study. *J Hematopathol*. 2022;15(4):203-213. doi:10.1007/s12308-022-00523-6

Comparison of the perinatal outcomes and antibiotic susceptibility patterns of acute cystitis associated with *Escherichia coli* and *Klebsiella pneumoniae* infections during the second trimester of pregnancy

 Muradiye Yıldırım,  Şevki Çelen

Department of Perinatology, Ankara Etlik City Hospital, Ankara, Türkiye

Cite this article as: Yıldırım M, Çelen Ş. Comparison of the perinatal outcomes and antibiotic susceptibility patterns of acute cystitis associated with *Escherichia coli* and *Klebsiella pneumoniae* infections during the second trimester of pregnancy. *J Health Sci Med.* 2025;8(6):1118-1123.

Received: 04.09.2025

Accepted: 10.10.2025

Published: 25.10.2025

ABSTRACT

Aims: To investigate the disease-causing potential of the most common uropathogens, *Escherichia coli* (*E. coli*) and *Klebsiella pneumoniae* (*K. pneumoniae*), as well as the effects of treatment alternatives on maternal and neonatal health, by analysing clinical outcomes and microbiological susceptibility patterns.

Methods: The retrospective cohort study included women diagnosed with acute cystitis between 14 and 28 weeks of gestation, who subsequently delivered. Those with chronic kidney disease, hypertension, diabetes mellitus, steroid therapy or fetal structural or chromosomal disorders were excluded. Obstetric, clinical and laboratory characteristics were recorded. Information on patients with recurrent episodes, progression of acute pyelonephritis, and outcomes of pregnancy and neonates was also recorded. The findings were analysed, and a p value of less than 0.05 was considered significant.

Results: A total of 276 pregnant women were diagnosed with acute cystitis and met the study criteria of urine culture positivity. *E. coli* growth was detected in 148 women (53.7%), *K. pneumoniae* in 42 women (15.2%). Patients in the *E. coli* group were younger and had a lower body-mass index than those in the *K. pneumoniae* group ($p < 0.001$). Laboratory values for white blood cells and C-reactive protein were higher in *E. coli* infections while hematocrit was lower ($p < 0.001$). Among those with *E. coli* growth, the extended-spectrum beta-lactamase (ESBL) positivity rate was 11.1% (n: 16), compared to 14.3% (n: 6) among those with *Klebsiella*, however no statistically significant difference was observed. Recurrence of infection and acute pyelonephritis progression rates were comparable in both groups. There was no difference in birth weeks, birth weight, preterm birth, premature rupture of the membranes, fetal growth restriction, preeclampsia, admission to the neonatal intensive care unit, neonatal infection or Apgar scores. The highest resistance rates in *E. coli* pathogens were found to be to ampicillin and amoxicillin-clavulanic acid, whereas in *K. pneumoniae* pathogens, resistance was found to be highest to amoxicillin-clavulanic acid and ceftriaxone. Extended-spectrum beta-lactamase positivity (ESBL)-positive *E. coli* strains were most susceptible to carbapenems, nitrofurantoin, piperacillin-tazobactam and fosfomycin. Similarly, ESBL-positive *K. pneumoniae* strains were most susceptible to carbapenems, piperacillin-tazobactam and fosfomycin.

Conclusion: The gold standard for identifying pathogens in urinary tract infections during pregnancy is remains urine combined culture with antibiotic susceptibility testing. Early screening, accurate diagnosis, and timely treatment of urinary tract infections are essential to safeguard both maternal and neonatal health. In obstetric practice, the judicious use of targeted antibiotics-appropriate in both spectrum and duration-is critical to minimize the development of antimicrobial resistance.

Keywords: *Escherichia coli*, *Klebsiella pneumoniae*, uropathogenes, acute cystitis, pregnancy, extended-spectrum beta-lactamase

INTRODUCTION

Urinary tract infections (UTI) are one of the most common infections during pregnancy. The incidence of UTIs is estimated at 8-10% and, by definition, they are regarded as complicated infections that require a tailored diagnostic approach and treatment.¹ Physiological, anatomical, hormonal and hemodynamic changes occur in the urinary system during pregnancy. Risk factors during pregnancy include ureteral and pelvic/lyceal dilatation, which is caused by

progesterone-induced smooth muscle relaxation; urine stasis, due to decreased bladder tone; increased intravesical pressure, induced by the growing uterus; vesicourethral reflux; elevated urine pH; and physiological glucosuria.² Although it can occur at any stage of pregnancy, it is most prevalent during the second trimester. UTIs during pregnancy present as asymptomatic bacteriuria (ASB), acute cystitis (AC) and acute pyelonephritis (APN). The new-onset of lower urinary tract symptoms

(dysuria, urgency, suprapubic pain) without systemic symptoms is known as UTI type AC. It is vital to diagnose and treat as quickly as possible to prevent complications for both the maternal and neonatal health.³ Untreated acute cystitis has been associated with an exacerbated risk of recurrence, pyelonephritis, sepsis, preterm birth, chorioamnionitis, low birth weight (LBW), preeclampsia and stillbirth.³⁻⁵

The two most common causes of urinary tract infections in pregnant women are *Escherichia coli* (*E. coli*) and *Klebsiella pneumoniae* (*K. pneumoniae*), which are both members of the gram-negative *Enterobacteriaceae* family of bacteria. *K. pneumoniae* is an opportunistic pathogen that has become increasingly common in hospital-acquired infections among people with compromised immune systems in recent years. The increasing prevalence of antibiotic-resistant bacteria and the production of extended-spectrum beta-lactamase (ESBL) reduces the options for antimicrobial agent during pregnancy. This can lead to the use of broader-spectrum antibiotics, longer hospital stays and treatment failure. There are limited studies in the literature regarding the length of hospital stay, increased morbidity, antibiotic resistance, financial burden and adverse perinatal outcomes associated with UTIs caused by different pathogens during pregnancy.⁵⁻⁹

This study aims to investigate the disease-causing potential, treatment alternatives and effects on maternal and neonatal health of the two most common pathogens by analysing clinical outcomes and microbiological susceptibility patterns. The findings will contribute to the more effective management of UTIs in pregnant women and help to ensure more informed empirical antibiotic choices are made.

METHODS

Study Population

The retrospective cohort study included patients diagnosed with acute cystitis between 14 and 28 weeks of gestation, who were treated as outpatients or in-patients. Data were obtained by checking the files of the Ankara Etlik City Hospital's Maternity Hospital between February 2023 and February 2025. The study was approved by the Ankara Etlik City Hospital Scientific Researches Evaluation and Ethics Committee (Date: 10.06.2025, Decision No: AEŞH-BADEK2-2025-131). The study followed the rules of the Declaration of Helsinki. Demographic and obstetric characteristics, antimicrobial susceptibility and resistance profiles, laboratory parameters, treatment regimens, gestational age, and obstetric and perinatal outcomes were examined retrospectively.

The study included low-risk pregnant women aged 18-45 years who were diagnosed with acute cystitis during the second trimester of pregnancy and subsequently gave birth in hospital. The presence of symptoms such as pain, frequent urination, urgency, and bacterial growth in the midstream urine culture sample (10^5 or more) were considered diagnostic.⁴ Recurrent acute cystitis was defined as experiencing at least two episodes of acute cystitis during pregnancy.¹⁰ Patients who were diagnosed with pyelonephritis or urosepsis during pregnancy, who did not have a urine culture taken before receiving antibiotic treatment, or who had chronic kidney disease or

were undergoing immunosuppressive therapy were excluded. Those with hypertension, diabetes mellitus, other chronic diseases, or fetal structural and chromosomal anomalies were excluded. Additionally, pregnant women with structural or functional urinary diseases, or who had been diagnosed with ASB, APN, were excluded from the study. The following were recorded; the patients' sociodemographic and obstetric characteristics at the week of AC diagnosis; age; number of pregnancies; parity; body-mass index (BMI); smoking habits; dipstick urine test results; the type of uropathogen identified in culture; the presence of ESBL production; and the results of white blood cell (WBC) ($10^9/L$), hematocrit (%) and C-reactive protein (CRP) (mg/L) tests. The presence of pyuria was determined by the presence of 5 or more leukocyte esterase on microscopic examination. Recurrent attacks of AC and those complicated by APN were determined.

The pregnant women were given oral fosfomycin, amoxicillin-clavulanic acid or nitrofurantoin after being diagnosed with AC, based on the results of their antibiograms. Those receiving parenteral treatment with ampicillin-sulbactam, ceftriaxone, piperacillin-tazobactam or carbapenem were hospitalised. Treatment lasted an average of three to seven days. Patients experiencing recurrent UTIs received treatment for 10-14 days.

Gestational age at diagnosis (week), birth weeks, birth weight (gram), mean fetal heart rate during labour (beat per minute), preterm birth, premature rupture of membranes (PROM), Fetal growth restriction (FGR), presence of preeclampsia, newborn 1- and 5-minute Apgar scores, neonatal intensive care unit (NICU) admission, and presence of neonatal infection were also questioned.

Statistical Analysis

The power of the study was determined using G Power software (version 3.1; Franz Faul, University, Kiel, Germany).¹¹ When the effect size was set to 0.5, the p-value to 0.05 and the statistical power to 80%, the minimum sample size each group was calculated to be 42. All statistical analyses were performed using IBM Statistical Package for the Social Sciences Statistics 27. Descriptive data are presented as mean $\pm$ standard deviation and percentage. Normality was determined using the Kolmogorov-Smirnov test. Data that were normally distributed were presented as mean $\pm$ SD. Parametric data were compared using an independent t-test. Data that were not normally distributed were presented as the median $\pm$ the 25th and 75th percentiles. Comparisons were made using the Mann-Whitney U test. Chi-square and Fisher's exact tests were used to compare categorical data. Statistical significance was assumed at an alpha level of 0.05.

RESULTS

A total of 276 pregnant women were diagnosed with acute cystitis and met the study criteria of urine culture positivity. *E. coli* growth was detected in 148 women (53.7%), *K. pneumoniae* in 42 women (15.2%), Streptococcus species in 25 women (9.1%), Proteus species in 19 women (6.9%), Staphylococcus species in 15 women (5.4%), Enterobacteriaceae species in 12 women (4.3%), Pseudomonas species 10 women (3.6%), and

other bacteria species in 5 women (1.8%). The mean age of the patients was 30.2 years, with mean gravida and parity of 3 and 1 respectively. Mean BMI was 27.5 kg/cm², and the gestational week of AC diagnosis was calculated as 22 weeks. Sixteen patients (6%) were smokers. Nitrite positivity in spot urine was detected in 100 patients (36.2%), pyuria in 164 patients (59.4%). ESBL-positive uropathogens were found in 29 cases (10.5%). Laboratory evaluation revealed a mean WBC count of 10.1 x 10⁹/L, a mean hematocrit value of 38.9%, and a mean CRP value of 9.5 mg/L. Nine women had recurrent cystitis attacks and seven women developed APN. The mean birth week of the patients was 38 weeks, the mean fetal heart rate during labour was 137 beats per minute, the average birth weight was 3.099 grams, and the mean Apgar scores were 7 at one minute and 9 at five minutes. Complications were categorised as preterm birth in 27 patients (9.8%), PROM in thirteen patients (4.7%), FGR in seventeen patients (6.2%), and preeclampsia in ten patients (3.6%). Thirty four newborns were admitted to the NICU, seventeen of whom (6.1%) developed a neonatal infection (Table 1).

Table 1. The sociodemographic, laboratory and clinical characteristics of all urinary tract infection patients

Variable (n=276)	
Age (years)	30.2±4.8
Gravidity	3±2
Parity	1±1
BMI (kg/cm ²)	27.5±4.1
Smoking	16 (6%)
Extended-spectrum beta-laktamase positivity	29 (10.5%)
Gestational age (weeks)	22±4
Nitrite positivity	100 (36.2%)
Pyuria	164 (59.4%)
White blood cell count (10 ⁹ /L)	10.1±3.7
Hematocrit (%)	38.9±4.3
C-reactive protein (mg/L)	9.5±4.8
Recurrent infections	9 (3.3%)
Progression of acute pyelonephritis	7 (2.5%)
Birth weeks	38±2
Preterm birth	27 (9.8%)
Premature rupture of membranes	13 (4.7%)
Fetal growth restriction	17 (6.2%)
Preeclampsia	10 (3.6%)
Birth weight (gram)	3099±453
Fetal heart rate (beat per minute)	137±6
Admission of neonatal intensive care unit	34 (12.3%)
Neonatal infection	17 (6.1%)
Apgar 1. minute	7±1
Apgar 5. minute	9±1

Values are presented as mean±standard deviation or as counts (percentage), BMI: Body-mass index

A comparison was made between the *E. coli* and *K. pneumoniae* groups. Patients in the *E. coli* group were younger and had a

lower BMI than those in the *K. pneumoniae* group ($p<0.001$). The rate of pyuria detection in the spot urine test was 90% in the *E. coli* group and 21.4% in the *Klebsiella* group ($p<0.001$). Nitrite positivity rates were similar. Laboratory values for WBC and CRP were higher in *E. coli* infection and Hct was lower ($p<0.001$). Among those with *E. coli* growth, the ESBL positivity rate was 11.1% (n: 16), compared to 14.3% (n: 6) among those with *Klebsiella*, though no statistically significant difference was observed. The recurrence of infection and the rate of APN progression were comparable in both groups. No difference was found in the other compared parameters in terms of clinical and perinatal adverse outcomes (Table 2). Data from ESBL-positive and negative isolates were compared (Table 3). WBC and CRP values were significantly higher in the ESBL-positive group than in the other groups. No differences were otherwise observed in laboratory or perinatal outcomes. While recurrent infection was more prevalent in ESBL-positive isolates, this was not reflected in a statistically significant difference.

Table 2. Obstetric, clinical and laboratory results according to uropathogens

	<i>K. pneumoniae</i> (n=42)	<i>E. coli</i> (n=148)	p-value
Age (years)	33.3±4.8	28.8±4.1	<.01*
Gravidity	3±2	2±1	.76*
Parity	1±1	1±1	.91*
Body-mass index (kg/cm ²)	29.4±3.5	26.8±3.8	.01*
Smoking	2 (4.8%)	12 (8%)	.52†
ESBL positivity	6 (14.3%)	16 (11.1%)	.61†
Gestational age (weeks)	22±4	22±5	.75*
Nitrite positivity	17 (40.5%)	62 (42.2%)	.85†
Pyuria	9 (21.4%)	133 (90%)	<.01†
White blood cell count (10 ⁹ /L)	7.8±3	11.2±3.5	<.01*
Hematocrit (%)	40.2±2.5	38.8±4.9	.01*
C-reactive protein (mg/L)	6.6±4.4	10.8±4.5	<.01*
Recurrent infections	1 (2.4%)	7 (4.7%)	.98††
Progression of acute pyelonephritis	0 (0%)	6 (4%)	.17††
Birth weeks	38±2	38±2	.99*
Preterm birth	3 (7.1%)	19 (12.8%)	.39††
Premature rupture of membranes	2 (4.8%)	8 (5.4%)	.99††
Fetal growth restriction	3 (7.1%)	10 (6.8%)	.99††
Preeclampsia	2 (4.8%)	5 (3.4%)	.65††
Birth weight (gram)	3030±389	3132±477	.23*
Fetal heart rate (beat per minute)	137±5	137±7	.82*
Admission of neonatal intensive care unit	8 (19.5%)	18 (12.2%)	.29††
Neonatal infection	4 (9.5%)	9 (6.3%)	.47††
Apgar 1. minute	7±1	7±1	.88*
Apgar 5. minute	8±2	8±1	.34*

*Independent T test, †Chi-square test, †† Fisher's exact test
Values are presented as mean±standard deviation or as counts (percentage)
K. pneumoniae: *Klebsiella pneumoniae*, *E. coli*: *Escherichia coli*, ESBL: Extended-spectrum beta-laktamase

Table 3. Comparison of obstetric, laboratory and clinical results of uropathogens according to the presence of extended-spectrum beta-laktamase

	ESBL (+) (n=22)	ESBL (-) (n=168)	p-value
Age (years)	31 (28-35)	30 (27-33)	.19\$
Gravidity	3 (1-3)	2 (1-3)	.69\$
Parity	2 (0-2)	1 (0-2)	.55\$
Body-mass index (kg/cm ²)	28.7 (26.1-30.9)	26.8 (24.5-30)	.11\$
Smoking	2 (9%)	12 (7.1%)	.92††
Gestational age (weeks)	21 (17-25)	22 (18-26)	.30\$
Nitrite positivity	16 (73%)	104 (61.9%)	.42†
Pyuria	6 (27.2%)	47 (27.9%)	.87††
White blood cell count (10 ⁹ /L)	14.6 (13.4-16.6)	9.5 (6.8-11.6)	<.01\$
Hematocrit (%)	39.3 (38.2-40.7)	39.9 (37-41.9)	.53\$
C-reactive protein (mg/L)	14.4 (12.9-17.2)	8.4 (5-12.2)	<.01\$
Recurrent infections	3 (13.6%)	5 (3%)	.11††
Progression of acute pyelonephritis	1 (6.3%)	4 (2.4%)	.41††
Birth weeks	38 (36-40)	38 (38-39)	.57\$
Preterm birth	5 (22.7%)	21 (12.5%)	.39††
Premature rupture of membranes	1 (4.5%)	9 (5.4%)	.92††
Fetal growth restriction	1 (6.3%)	12 (7.1%)	.92††
Preeclampsia	0 (0%)	7 (4.3%)	.40††
Birth weight (gram)	3035 (2560-3360)	3155 (2860-3470)	.36\$
Fetal heart rate (beat per minute)	138 (134-141)	136 (134-139)	.51\$
Admission of neonatal intensive care unit	5 (22.7%)	22 (13.1%)	.25††
Neonatal infection	1 (6.3%)	12 (7.1%)	.90††
Apgar 1.minute	7 (7-8)	7 (7-8)	.69\$
Apgar 5.minute	9 (8-9)	9 (8-9)	.78\$

§Man-Whitney U test, †Chi-square test, ††Fisher's exact test
 Values are presented as median (25th-75th percentile) or as counts (percentage)
 ESBL: Extended-spectrum beta-laktamase

The highest resistance rates in *E. coli* pathogens were found to be to ampicillin and amoxicillin-clavulanic acid, whereas in *K. pneumoniae* pathogens, they were found to be to amoxicillin-clavulanic acid and ceftriaxone. ESBL-positive *E. coli* strains were most susceptible to carbapenems, nitrofurantoin, piperacillin-tazobactam and fosfomycin. Similarly, ESBL-positive *K. pneumoniae* strains were most susceptible to carbapenems, piperacillin-tazobactam and fosfomycin.

DISCUSSION

The current manuscript investigates the antibiotic susceptibility profiles of *E. coli* and *K. pneumoniae* pathogens and their impact on pregnancy outcomes in AC, one of the most common infections encountered in pregnant women. The *E. coli* pathogen was found to pose a risk to young, lean women. More laboratory findings were found to be caused by *E. coli* cystitis. Increased WBC and CRP levels, decreased hematocrit and pyuria on dipstick testing are more common in AC caused by *E. coli*. It can be assumed that infections caused by *E. coli* require closer monitoring to detect recurrence or

progression to pyelonephritis. The two pathogens were found to have comparable ESBL rates. Adverse pregnancy outcomes and neonatal complications were found to be similar.

AC is characterised by symptoms of the lower urinary tract and a positive urine culture, but no systemic symptoms. One to two percent of all UTIs are accounted for by this condition, as are 25 percent of symptomatic UTIs during pregnancy.¹² Most pregnant women presenting with AC in the second trimester will have had a negative urine culture screening test in the first trimester.¹ A positive urine culture result at any stage of pregnancy in the absence of urinary symptoms is known as ASB. Treatment is not required except in cases of pregnancy, urological procedures and transplant patients.^{2,6} Routine ASB treatment during pregnancy reduces the risk of pyelonephritis.^{3,13,14} However, the lower urinary tract symptoms observed in AC can be confused with typical pregnancy symptoms, in contrast to ASB. Therefore, a urinalysis is necessary. The diagnosis of AC is confirmed in the presence of new-onset symptomatic lower urinary tract complaints without systemic findings and colony growth of 10⁵ or more. However, some authors recommend treatment with a colony count of 10² for a single organism in the presence of symptoms other than ASB.¹² In our study, WBC and CRP levels were significantly higher, and hematocrit values were lower, in the *E. coli* group. Because uropathogenic *E. coli* strains induce a strong inflammatory response, findings such as leukocytosis and elevated CRP are generally more pronounced in cases of pyelonephritis, according to the literature.^{5,15} However, the current literature lacks clear, comparative evidence regarding the frequency or severity of anemia by causative agent. In our study, pyuria was detected in 90% of cases of *E. coli* cystitis. Although pyuria has a sensitivity of over 97% for UTIs, its specificity is lower as it can be contaminated from the vulva and vagina. The presence of nitrite is a more specific indicator. Therefore, if neither nitrite nor pyuria are present in a dipstick test, a UTI can be ruled out.¹² Early and appropriate treatment, guided by susceptibility testing, is essential in reducing complications and recurrence rates. Delayed or inadequate treatment, particularly in the presence of resistant strains, can lead to persistent infection and an increased risk of maternal and neonatal morbidity.¹⁶

The study revealed no significant differences in adverse pregnancy outcomes between women who underwent AC with two agents. Additionally, the recurrence rates and progression to pyelonephritis of AC caused by both uropathogens were found to be similar. In our study, all cases complicated by APN were recurrent *E. coli* infections. Almost all cases of PROM occurred in pregnant women with a history of recurrent infection or pyelonephritis. This is consistent with the inflammatory pathophysiology of rupture of membranes.¹⁷ Symptomatic lower urinary tract infections during pregnancy increase the risk of spontaneous preterm birth, LBW and postpartum endometritis and mastitis, even in low-risk pregnancies.¹⁸ The current study found that the frequency of adverse perinatal outcomes in the sample was similar to that in the general population.¹⁸ Our study may have focused on community-acquired and second-trimester-specific infections, which could explain this. A retrospective

study compared *Proteus*, *Klebsiella*, *Enterobacter*, *Citrobacter*, *Acinetobacter*, *Staphylococcus*, and *Raoultella* species, pathogens of ASB or AC, with *E. coli* in 313 pregnant women.⁷ No differences were found in the rates of pyelonephritis progression, length of hospital stay, preterm birth or LBW. The results of this study, in which the morbidities of these strains, which are considered more pathogenic outside pregnancy, were compared with *E. coli*, are consistent with those we obtained. This study is different from our current study in that it included UTIs during pregnancy at all trimesters. It also did not provide information on whether the pathogens were acquired in hospital or in the community.⁷ Another study that comparing the result of *E. coli* and group B Streptococcus lower UTIs found that the progression of pyelonephritis was higher with *E. coli*, but that obstetric complications were similar.⁵ In a study of a non-pregnant APN population, catheter-related infection, community-acquired infection and healthcare-associated APN were observed at a higher rate in *K. pneumoniae* than in *E. coli*. More severe APN was observed with *E. coli*.¹⁵

The present study, ESBL production was found to be higher in *K. pneumoniae* than in *E. coli*, although no statistically significant differences were observed. ESBL is an enzyme produced especially by *E. coli* and *K. pneumoniae* bacteria that limits treatment options by inactivating third and fourth generation cephalosporins and monobactams. The rise in antibiotic resistance is a significant problem, with varying levels across different societies.⁸ A study investigating resistance rates in urinary tract infections in pregnant women in our country found that ESBL production was 16.04 per cent in *E. coli* and 20.9 per cent in *K. pneumoniae*.⁹ Antibiotic sensitivity rates play a decisive role in the selection of prophylactic antibiotics for pregnant women. The current study found that *E. coli* pathogens with positive ESBL tests had no resistance to carbapenems or nitrofurantoin, and above 90% were susceptible to piperacillin-tazobactam. *K. pneumoniae* pathogens with positive ESBL tests exhibited nearly 100% susceptibility to carbapenems and piperacillin-tazobactam, followed by fosfomycin and nitrofurantoin. The treatment plans for our patients were impacted by these resistance profiles, making outpatient treatment challenging and necessitating inpatient care for infections treated with carbapenems and piperacillin-tazobactam. Prophylactic antibiotic therapy can be initiated following diagnosis, with subsequent treatment based on culture results and pathogen sensitivity. Amoxicillin and ampicillin regimens should be avoided if prophylactic antibiotic therapy is initiated due to increased resistance to *E. coli*.^{4,19} Nitrofurantoin is a common alternative for treating UTIs due to urinary tract pathogens becoming increasingly resistant to penicillin derivatives, and nitrofurantoin being effective against a broad range of Gram-negative and gram-positive organisms. Most animal and human studies have not found a link between first-trimester nitrofurantoin use and an elevated risk of congenital anomalies.^{4,20,21} However, the methodological differences in some studies that have found an association with hypoplastic left heart syndrome are being questioned.^{21,22} In conclusion, it has been reported that the drug does not increase the risk of malformations in the first trimester and that nitrofurantoin

and sulfonamides are suitable options in the absence of suitable alternatives.⁴ American College of Obstetrics and Gynecology accepts it as a first-line treatment option for lower UTIs in second trimester pregnancies.⁴ The use of this medication is contraindicated in patients with glucose-6-phosphate dehydrogenase deficiency, as it can result in pulmonary toxicity, hemolytic anaemia, and neonatal jaundice.^{4,23}

Limitations

The study's strengths include its focus on the second trimester, its specific focus on acute cystitis, and the fact that all infections were community-acquired. The limitations include the study's retrospective nature, its relatively small sample size, its single-centre design and the lack of molecular-level analysis. One of the limitations of our study is the imbalance in the number of samples between groups. Future research could investigate resistance mechanisms at a molecular level. Multicentre, prospective studies with larger sample sizes could also be planned. Strategies to reduce resistance or develop vaccines could be developed.

CONCLUSION

Improved maternal and neonatal outcomes are dependent on early diagnosis of UTI during pregnancy and treatment that is specific to the pathogen. In obstetric practice, the judicious use of targeted antibiotics-appropriate in both spectrum and duration-is critical to minimize the development of antimicrobial resistance.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Ankara Etlik City Hospital Scientific Researches Evaluation and Ethics Committee (Date: 10.06.2025, Decision No: AEŞH-BADEK2-2025-131).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Acknowledgements

I would like to express my special thanks to all healthcare staff who have worked with devotion.

REFERENCES

- Matuszkiewicz-Rowińska J, Małyszko J, Wieliczko M. Urinary tract infections in pregnancy: old and new unresolved diagnostic and therapeutic problems. *Arch Med Sci*. 2015;11(1):67-77. doi:10.5114/aoms.2013.39202
- Ansaldi Y, Martinez de Tejada Weber B. Urinary tract infections in pregnancy. *Clin Microbiol Infect*. 2023;29(10):1249-1253. doi:10.1016/j.cmi.2022.08.015
- Balachandran L, Jacob L, Al Awadhi R, et al. Urinary tract infection in pregnancy and its effects on maternal and perinatal outcome: a retrospective study. *Cureus*. 2022;14(1):e21500. doi:10.7759/cureus.21500
- Urinary tract infections in pregnant individuals. *Obstet Gynecol*. 2023;142(2):435-445. doi:10.1097/AOG.0000000000005269
- Sachdeva S, Rosett HA, Krischak MK, et al. Urinary tract infection and progression to pyelonephritis: group B *Streptococcus* versus *E. coli*. *AJP Rep*. 2024;14(1):e80-e84. doi:10.1055/s-0044-1779031
- Abde M, Weis N, Kjærbye-Thygesen A, Moseholm E. Association between asymptomatic bacteriuria in pregnancy and adverse pregnancy- and births outcomes. A systematic review. *Eur J Obstet Gynecol Reprod Biol*. 2024;302:116-124. doi:10.1016/j.ejogrb.2024.09.001
- Rosett HA, Krischak MK, Sachdeva S, et al. Lower urinary pathogens: do more pathogenic bacteria increase the risk of pyelonephritis? *Am J Perinatol*. 2022;39(5):473-478. doi:10.1055/s-0040-1717093
- Rosana Y, Ocviyanti D, Halim M, et al. Urinary Tract Infections among Indonesian Pregnant Women and Its Susceptibility Pattern. *Infect Dis Obstet Gynecol*. 2020;2020:9681632. doi:10.1155/2020/9681632
- Aydoğmuş S, Kılıç EK. Determination of antibiotic resistance rates of *Escherichia coli* and *Klebsiella pneumoniae* isolates, which are the causative agents of urinary tract infection in pregnant women. *Anat Curr Med J*. 2023;5(2):97-101. doi:10.38053/acmj.1240882
- Epp A, Larochelle A. No. 250-recurrent urinary tract infection. *J Obstet Gynaecol Can*. 2017;39(10):e422-e431. doi:10.1016/j.jogc.2017.08.017
- Faul F, Erdfelder E, Lang AG, Buchner A. G*power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods*. 2007;39(2):175-191. doi:10.3758/bf03193146
- Chu CM, Lowder JL. Diagnosis and treatment of urinary tract infections across age groups. *Am J Obstet Gynecol*. 2018;219(1):40-51. doi:10.1016/j.ajog.2017.12.231
- Nicolle LE, Gupta K, Bradley SF, et al. Clinical practice guideline for the management of asymptomatic bacteriuria: 2019 Update by the Infectious Diseases Society of America. *Clin Infect Dis*. 2019;68(10):1611-1615. doi:10.1093/cid/ciz021
- Gilstrap LC 3rd, Ramin SM. Urinary tract infections during pregnancy. *Obstet Gynecol Clin North Am*. 2001;28(3):581-591. doi:10.1016/s0889-8545(05)70219-9
- Hyun M, Lee JY, Kim HA, Ryu SY. Comparison of *Escherichia coli* and *Klebsiella pneumoniae* acute pyelonephritis in Korean patients. *Infect Chemother*. 2019;51(2):130-141. doi:10.3947/ic.2019.51.2.130
- Kalinderi K, Delkos D, Kalinderis M, Athanasiadis A, Kalogiannidis I. Urinary tract infection during pregnancy: current concepts on a common multifaceted problem. *J Obstet Gynaecol*. 2018;38(4):448-453. doi:10.1080/01443615.2017.1370579
- Dayanan R, Bayraktar B, Karabay G, et al. Prognostic evaluation of albumin-associated inflammatory indices (LAR, NAR, MAR, FAR, SIRI/Alb, SIRI/Alb, NPAR, and CAR) in preterm premature rupture of membranes (PPROM) pregnancies. *BMC Pregnancy Childbirth*. 2025;25(1):759. doi:10.1186/s12884-025-07877-6
- Werter DE, Schneeberger C, Mol BWJ, et al. The risk of preterm birth in low risk pregnant women with urinary tract infections. *Am J Perinatol*. 2023;40(14):1558-1566. doi:10.1055/s-0041-1739289
- Ribeiro-do-Valle CC, Bonet M, Brizuela V, et al. Aetiology and use of antibiotics in pregnancy-related infections: results of the WHO Global Maternal Sepsis Study (GLOSS), 1-week inception cohort. *Ann Clin Microbiol Antimicrob*. 2024;23(1):21. doi:10.1186/s12941-024-00681-8
- Goldberg O, Koren G, Landau D, Lunenfeld E, Matok I, Levy A. Exposure to nitrofurantoin during the first trimester of pregnancy and the risk for major malformations. *J Clin Pharmacol*. 2013;53(9):991-995. doi:10.1002/jcph.139
- Goldberg O, Moretti M, Levy A, Koren G. Exposure to nitrofurantoin during early pregnancy and congenital malformations: a systematic review and meta-analysis. *J Obstet Gynaecol Can*. 2015;37(2):150-156. doi:10.1016/S1701-2163(15)30337-6
- Osmundson SS, Nickel KB, Shortreed SM, et al. First-trimester antibiotic use for urinary tract infection and risk of congenital malformations. *JAMA Netw Open*. 2025;8(7):e2519544. doi:10.1001/jamanetworkopen.2025.19544
- Nordeng H, Lupattelli A, Romøren M, Koren G. Neonatal outcomes after gestational exposure to nitrofurantoin. *Obstet Gynecol*. 2013;121(2 Pt 1):306-313. doi:10.1097/AOG.0b013e31827c5f88

Role of first trimester inflammatory markers in predicting adverse obstetric and neonatal outcomes in maternal obesity

 Betül Akgün Aktaş¹,  Emre Gençslan²,  Ezgi Başaran³,  Ayşe Altındış Bal²,
 Burcu Bozkurt Özdağ²,  Fatma Doğa Öcal²,  Dilek Şahin²

¹Department of Perinatology, Uşak Training and Research Hospital, Uşak, Türkiye

²Department of Perinatology, Ankara Bilkent City Hospital, Ankara, Türkiye

³Department of Obstetrics and Gynecology, Kayseri City Hospital, Kayseri, Türkiye

Cite this article as: Akgün Aktaş B, Gençslan E, Başaran E, et al. Role of first trimester inflammatory markers in predicting adverse obstetric and neonatal outcomes in maternal obesity. *J Health Sci Med.* 2025;8(6):1124-1128.

Received: 14.09.2025

Accepted: 11.10.2025

Published: 25.10.2025

ABSTRACT

Aims: To compare first trimester serum inflammatory markers in obese pregnant women with healthy pregnant women and to investigate the predictive value of these markers for pregnancy and neonatal outcomes

Methods: This prospective observational study included 46 pregnant women with maternal obesity and 60 healthy pregnant women without obesity. The demographic characteristics, obstetric and neonatal outcomes of the participants were evaluated, and the neutrophil-lymphocyte ratio (NLR), systemic inflammation index (SII: $NLR \times platelet$), systemic inflammation response index (SIRI: $NLR \times monocyte$), aggregate index of systemic inflammation (AISI: $NLR \times platelet \text{ count} \times monocyte$), and PLR (platelet-lymphocyte ratio) were calculated. To evaluate the predictive value for neonatal intensive care unit (NICU) admission and adverse obstetric outcomes, receiver operating characteristic curves were generated and the area under the curve values were calculated.

Results: The demographic characteristics of both groups were similar. In the study group, adverse obstetric outcomes, and NICU requirement were increased. In the maternal obese group, inflammatory markers such as SII, SIRI, AISI, and NLR were significantly elevated. PLR was also increased in the study group. In ROC analysis, it was found that the levels of inflammatory markers to predict adverse obstetric and neonatal outcomes in obese pregnant women were statistically significant.

Conclusion: SII, SIRI, AISI, PLR, and NLR may be helpful in predicting neonatal and obstetric outcomes in maternal obesity.

Keywords: Obesity, pregnancy, systemic inflammation index, aggregate index of systemic inflammation

INTRODUCTION

Obesity continues to be a prevalent health problem worldwide. According to data from the World Health Organization, one in eight people live with obesity.¹ Obesity, particularly observed in women of reproductive age, leads to short- and long-term adverse obstetric and neonatal outcomes.² The inappropriate environment caused by the metabolic syndrome accompanying obesity-including hyperinsulinemia, chronic inflammation, increased oxidative stress, endothelial dysfunction, and vascular involvement-along with placental abnormalities, increases morbidity and mortality in both mothers and offspring. In obese mothers, increases may be seen in preeclampsia, diabetes, fetal distress, pregnancy loss, and the need for neonatal intensive care.³⁻⁵

In recent years, systemic inflammatory markers, which are indicators of inflammation and can be easily calculated from a complete blood count, have been shown to provide insight into disease prognosis during inflammatory processes. In a systemic inflammatory process, an increase in neutrophil count and a decrease in lymphocyte count were first defined as the neutrophil-to-lymphocyte ratio (NLR) in 2001 and

used as a prognostic value.⁶ In previous studies, with the addition of monocyte and platelet counts to neutrophil and lymphocyte counts, the Systemic Inflammation Index (SII), Systemic Inflammation Response Index (SIRI), and Systemic Inflammation Aggregate Index (AISI) were developed. These inflammatory indices have been used in the course and prognostication of many diseases, including heart disease, cancer, and chronic illnesses.^{7,8}

Obesity has also become a serious health problem in pregnant women and creates a hyperinflammatory environment in both the mother and fetus. In this study, we aimed to compare first trimester serum inflammatory markers in obese pregnant women with those in healthy pregnant women and to investigate the predictive value of these markers for pregnancy and neonatal outcomes.

METHODS

This study is a prospective observational study, and ethical approval was obtained from the Ankara Bilkent City Hospital

Corresponding Author: Betül Akgün Aktaş, drbetul07@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

Ethics Committee (Date: 13.11.2024, Decision No: TABED 2-24-642). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki. Written and oral informed consent was obtained from all participants. A total of 106 pregnant women, aged between 18 and 45 years, whose pregnancies were followed and deliveries occurred at the hospital between the years 2024-2025, were included in the study. The obese group consisted of 46 pregnant women with a body-mass index (BMI) of 30 kg/m² and above, who voluntarily agreed to participate in the study. The control group consisted of 60 pregnant women with a BMI below 30 kg/m², who had similar age and baseline characteristics to the study group and were randomly selected among those who wished to participate. Exclusion criteria included multiple pregnancy, chronic hypertension, pregestational diabetes, smoking, presence of any inflammatory disease, and a history of cancer.

Characteristics such as age, nulliparity, BMI, and neonatal outcomes of the participants were included in the study. According to the national guideline, the assessment of pregnancy up to the 14th gestational week is referred to as the first visit, during which a routine complete blood count is evaluated. From the participants' complete blood count results, the NLR, Systemic Inflammation Index (SII: NLR * platelet count), Systemic Inflammation Response Index (SIRI: NLR * monocyte count), systemic inflammation aggregate index (AISI: NLR * platelet count * monocyte count), PLR (platelet-to-lymphocyte ratio), and monocyte-to-lymphocyte ratio (MLR) were calculated.

Adverse obstetric outcomes included preterm labor, preterm premature rupture of membranes, hypertensive disorders of pregnancy, gestational diabetes, deep vein embolism, macrosomia (birth weight >4000 g), and fetal distress. Neonatal birth weight, APGAR score, and the need for neonatal intensive care unit (NICU) admission were recorded as separate outcomes. Inflammatory indices addressed in the study were analyzed to predict adverse obstetric history and NICU admission.

Statistical Analysis

Data analysis was performed using IBM SPSS Statistics 26.0 (IBM Corp., Armonk, NY, USA). The normality of the variables was tested using both the Shapiro-Wilk and Kolmogorov-Smirnov tests. Data were presented as median and interquartile range (IQR), number (n), and percentage (%) and were analyzed using the Mann-Whitney U test for non-parametric variables. A p-value <0.05 was considered statistically significant. Receiver operating characteristic (ROC) curves were generated to assess the predictive value of NICU admission and adverse obstetric outcomes, and the area under the curve (AUC) values were calculated using standard methods. The data were presented as AUC ROC (95% CI). Optimal cut-off values for the inflammatory indices were calculated using the Youden index. In a similar study, estimation of sample size required at least 26 participants per group, assuming a 5% margin of error and 90% power.⁹

RESULTS

Table 1 presents the baseline characteristics of the participants. Maternal age, nulliparity, and the gestational week at which

the complete blood count was evaluated were similar between the two groups. In the obese group, 21 participants were nulliparous, compared with 25 in the healthy control group. The gestational weeks at delivery and birth weight were found to be significantly lower in the obese group compared to the healthy controls. While the APGAR scores were similar between the two groups, the obese group had higher rates of emergency cesarean section, adverse obstetric outcomes, and need for NICU admission. In the subgroup analysis according to parity and emergency cesarean section rates, the highest rate of emergency cesarean section was observed in the obese nulliparous group (n=11, 52.4%), whereas the lowest rates were found in the obese multiparous group (n=3, 12%) and the multiparous control group (n=5, 14.3%) (p=0.030).

Table 1. Baseline characteristics of participants

	Obese group (n=46)	Control group (n=60)	p-value
Maternal age (years)	30 (11)	27 (7)	0.053
Gestational age at scan (week)	9 (5)	9 (5.7)	0.411
Pregestasyonel BMI (kg/m ²)	33.5 (5.7)	26.6 (4.2)	<0.001
Nulliparity (n, %)	21 (45.6)	25 (41.6)	0.682
Gestational age at birth (week)	38 (3)	39 (2)	0.001
Birth weight (gram)	3210 (715)	3290 (495)	0.001
Emergency cesarean section (n, %)	14 (30.4)	15 (25)	0.006
APGAR score at 1 st minute	7 (1)	8 (1)	0.101
APGAR score at 5 th minute	9 (0)	9 (0)	0.208
NICU admission (n, %)	6 (13)	1 (0.01)	0.019
Adverse perinatal outcomes (n, %)	17 (36.9)	11 (18.3)	0.044

Note: Data given as median (interquartile range), number, percentile (%). p-values <0.05 were considered statistically significant.
BMI: Body-mass index, APGAR: Appearance, pulse, grimace, activity and respiration, NICU: Neonatal intensive care unit

In the obese group, six patients experienced fetal distress, two had hypertension, three had preterm labor, two had macrosomia, and one patient each had preterm premature rupture of membranes, uterine atony, deep vein thrombosis, and gestational diabetes. In the control group, fetal distress occurred in six patients, hypertension in two, and macrosomia in three. However, when each adverse obstetric outcome was analyzed separately, no statistically significant difference was observed between the groups.

Table 2 presents the first-trimester blood cell counts and inflammatory parameters for both groups. There was no difference in hemoglobin levels or lymphocyte counts between the obese and control groups. However, white blood cell count, neutrophil and monocyte counts, and platelet count-recognized as inflammatory markers-were significantly higher in the obese group (p=0.001; p<0.001; p=0.032; p=0.010, respectively). Additionally, the maternal obese group showed significantly higher levels of SII, SIRI, AISI, and NLR. The PLR was 138.8 (55.2) in the study group and 131 (48.7) in the control group (p=0.052). When participants were analyzed by parity, multiparous obese women (n=25) had significantly higher NLR (3.4 (1.7) vs. 2.8 (1); p=0.002), SIRI (1.4 (0.8) vs. 1.0 (0.7); p=0.001), AISI (353.3 (255.7) vs. 228.1 (165); p=0.001), and SII (840 (373.3) vs. 663.3 (230); p<0.001) compared to

multiparous non-obese women (n=35). In nulliparous obese women (n=21), SII and AISI were significantly higher than in nulliparous non-obese women (n=25) (SII: 918.9 vs. 632.2; p=0.003 and AISI: 431.6 vs. 269.4; p=0.008, respectively).

Table 2. Blood cell counts and inflammatory parameters in the first trimester of pregnancy in both groups			
	Obese group (n=46)	Control group (n=60)	p-value
Hemoglobin (g/dl)	12.5 (0.7)	12.5 (1.4)	0.182
Platelet (10 ⁹ /L)	254 (77)	237.5 (66.2)	0.010
WBC (10 ⁹ /L)	9.2 (3.9)	7.6 (2.8)	0.001
Neutrophil (10 ⁹ /L)	7 (2.6)	4.9 (1.9)	<0.001
Lymphocyte (10 ⁹ /L)	1.9 (0.7)	1.8 (0.7)	0.559
Monocyte (10 ⁹ /L)	0.5 (0.1)	0.3 (0.1)	0.032
Eosinophil (10 ⁹ /L)	0.1 (0.05)	0.06 (0.1)	0.465
SII	903.9 (622)	610.5 (228.8)	<0.001
SIRI	1.83 (0.8)	0.90 (0.6)	0.001
AISI	473.1 (343.6)	206.6 (176.6)	<0.001
NLR	3.5 (2.1)	2.5 (0.9)	0.002
MLR	0.25 (0.04)	0.19 (0.08)	0.059
PLR	138.8 (55.2)	131 (48.7)	0.052
Note: Data given as median (interquartile range), number, percentile (%). p-values <0.05 were considered statistically significant. WBC: White blood cell, SII: Systemic Inflammation Index, SIRI: Systemic Inflammatory Response Index, AISI: Aggregated Index of systemic inflammation, NLR: Neutrophil-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio, PLR: Platelet-to- lymphocyte ratio			

Tables 3 and 4 analyze the predictive value of systemic inflammatory indices for adverse obstetric outcomes and NICU admission (**Figure 1, 2**). For predicting adverse obstetric outcomes, the optimal cut-off values were 723.8 for SII and 282.2 for AISI, respectively (sensitivity 67.3%, specificity 67.9%, AUC 0.712; p=0.002 vs. sensitivity 64.3%, specificity 62.5%, AUC 0.681; p=0.007). For predicting NICU admission, the optimal cut-off values were 804.9 for SII and 348 for AISI, respectively (sensitivity 71.4%, specificity 71.2%, AUC 0.824; p=0.005 vs. sensitivity 71.4%, specificity 71.2%, AUC 0.732; p=0.044). However, when adverse obstetric outcomes were analyzed by subgroups, these parameters did not retain statistical significance.

Table 3. The ROC curve analysis for the performance of first trimester inflammatory markers in predicting adverse obstetric outcomes						
Variable	Cut-off	AUC	p-value	95% CI	Sensitivity	Specificity
NLR	2.8	0.549	0.468	0.441-0.687	50%	50%
SII	723.8	0.712	0.002	0.596-0.827	67.3%	67.9%
SIRI	1.21	0.638	0.039	0.510-0.767	60.7%	60.7%
AISI	282.2	0.681	0.007	0.559-0.803	64.3%	62.5%
PLR	138.6	0.517	0.789	0.386-0.648	52.9%	52.4%
Note: p-values <0.05 were considered statistically significant. ROC: Receiver operating characteristic, AUC: Area under the curve, CI: Confidence interval, NLR: Neutrophil-to-lymphocyte ratio, SII: Systemic Inflammation Index, SIRI: Systemic Inflammatory Response Index, AISI: Aggregated Index of systemic inflammation, PLR: Platelet-to- lymphocyte ratio						

Table 4. The ROC curve analysis for the performance of first trimester inflammatory markers in predicting NICU admission						
Variable	Cut-off	AUC	p-value	95% CI	Sensitivity	Specificity
NLR	3.04	0.704	0.077	0.489-0.918	57.1%	61.6%
SII	804.9	0.824	0.005	0.707-0.941	71.4%	71.2%
SIRI	1.23	0.661	0.160	0.465-0.858	57.1%	56.2%
AISI	348	0.732	0.044	0.563-0.901	71.4%	71.2%
PLR	148	0.716	0.060	0.493-0.940	71.4%	67.1%
Note: p-values <0.05 were considered statistically significant. ROC: Receiver operating characteristic, NICU: Neonatal intensive care unit, AUC: Area under the curve, CI: Confidence interval, NLR: Neutrophil-to-lymphocyte ratio, SII: Systemic Inflammation Index, SIRI: Systemic Inflammatory Response Index, AISI: Aggregated Index of systemic inflammation, PLR: Platelet-to- lymphocyte ratio						

DISCUSSION

To the best of our knowledge, this is the first study to investigate the relationship between serum inflammatory parameters and obstetric outcomes in maternal obesity. In obesity, increased visceral adipose tissue and elevated free fatty acids lead to a shift from an optimal intrauterine environment to a proinflammatory milieu due to elevated cytokines such as TNF- α , IL-1, and IL-6, as well as reactive oxygen species, ultimately resulting in chronic inflammation.¹⁰ Furthermore, maternal obesity can cause oxidative stress damage in the placenta, altering DNA methylation and mRNA expression, thereby affecting the fetus during intrauterine life.¹¹ One study reported that the chronic low-grade inflammatory environment and fetal malprogramming induced by obesity negatively affected neurodevelopment in offspring of obese mothers, showing disruptions in microglia and synaptic structures.¹² Similarly, in a murine model, maternal obesity was shown to cause fetal cardiac remodeling and dysfunction.¹³ Beyond fetal anomalies, obesity has also been associated with adverse obstetric outcomes. In obese pregnant women, similar to our study, the rate of primary cesarean section due to fetal distress has been found to be increased.¹⁴ Particularly in pregnancies with a BMI over 40 kg/m², the risk of fetal acidosis at birth is elevated.¹⁵ Consistent with these findings, a higher rate of adverse obstetric outcomes in the obese group was observed in our study. Although we did not find statistically significant differences among the subgroups of adverse obstetric outcomes in this study, we believe that this may be due to the limited number of participants. A meta-analysis has shown an association between high maternal BMI and macrosomia, preterm birth, and neonatal asphyxia.¹⁶ Another recent meta-analysis reported increased rates of gestational diabetes and hypertension with preeclampsia in obese women.¹⁷

In a 7-year population-based study including over 20,000 adult participants, obesity was found to be positively associated with systemic inflammatory parameters.¹⁸ Similarly, a recent large-scale study demonstrated a positive correlation between visceral adipose tissue and the SII in adults.¹⁹ Increased visceral adipose tissue in obese individuals may contribute to a pro-inflammatory environment through elevated adipokines and proinflammatory mediators. Comparable results have been observed in obese children, where NLR, SIRI, SII, and PLR were higher in obese children diagnosed with metabolic

syndrome compared to obese children without metabolic syndrome.²⁰

Systemic inflammatory parameters have been used to predict various maternal diseases, adverse pregnancy outcomes, and complications such as preeclampsia, fetal growth restriction, preterm premature rupture of membranes, and diabetes. In pregnancies diagnosed with gestational diabetes, elevated first-trimester SIRI and SII values have been reported.²¹ In another study, high SII and AISI levels were found to be statistically significant in predicting gestational diabetes; however, no significant association was observed with adverse obstetric outcomes or NICU admissions.²² In patients undergoing cervical cerclage due to cervical insufficiency, high SII and SIRI were associated with unsuccessful cerclage and poor neonatal outcomes.²³ Similarly, in our study, elevated SII and AISI were statistically significantly associated with adverse obstetric outcomes and NICU admission. Additionally, we found that although within normal limits, platelet counts were higher in the obese group compared to the healthy controls. Recent studies suggest that platelets (PLT) play a role not only in hemostasis but also in inflammation.^{24,25} Upon PLT activation, soluble inflammatory mediators are released, and the migration and activation of cells involved in inflammation, such as monocytes, leukocytes, and macrophages, are triggered. In a study conducted on adults, PLT counts and SII were higher in the obese group, while no statistically significant difference was found in PLR.²⁶ Another study in children with obesity reported elevated levels of neutrophils, PLT, and PLR.²⁷ In a large study involving nearly 3,000 children and adolescents, a significant positive correlation was found between BMI and both SII and SIRI.²⁸

The results of studies examining maternal diseases are consistent with studies on systemic inflammatory parameters and disease course and prognosis. In pregnant women with familial mediterranean fever (FMF), first-trimester NLR, SII, and SIRI levels were higher compared to healthy pregnant women, and these inflammatory markers were also elevated in FMF patients with perinatal complications compared to those without complications.²⁹ Similar findings have been reported in previous studies involving pregnant women with systemic lupus erythematosus, thrombophilia, and ulcerative colitis.^{9,30,31} An interesting study on systemic inflammatory parameters found that SIRI, SII, and NLR were elevated in cesarean scar pregnancies, and high SII levels were significantly associated with predicting scar pregnancy in women with a history of cesarean delivery.³²

Limitations

While our study is valuable in that it includes first-trimester serum inflammatory parameters and evaluates pre-pregnancy BMI and birth outcomes, its limitations include the relatively small sample size and the lack of subgroup analysis for adverse obstetric outcomes. Further large-scale studies on this topic are needed to obtain more definitive results.

CONCLUSION

SII, SIRI, AISI, PLR, and NLR may be helpful in predicting neonatal and obstetric outcomes in maternal obesity.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Ankara Bilkent City Hospital Ethics Committee (Date: 13.11.2024, Decision No: TABED 2-24-642).

Informed Consent

All patients signed and free and informed consent form.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Acknowledgments

The authors would like to thank all healthcare professionals who are dedicated to combating obesity and maternal-fetal health disorders, and to protecting the well-being of mothers and infants. The data supporting the findings of this study are available from the corresponding author (BAA) upon reasonable request.

REFERENCES

1. World Health Organization (WHO) Accessed 7 May 2025 Available from: <https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight>
2. Denizli M, Capitano ML, Kua KL. Maternal obesity and the impact of associated early-life inflammation on long-term health of offspring. *Front Cell Infect Microbiol.* 2022;12:940937. doi:10.3389/fcimb.2022.940937
3. Catalano PM, Shankar K. Obesity and pregnancy: mechanisms of short term and long term adverse consequences for mother and child. *BMJ.* 2017;356:j1. doi:10.1136/bmj.j1
4. Marchi J, Berg M, Dencker A, Olander EK, Begley C. Risks associated with obesity in pregnancy, for the mother and baby: a systematic review of reviews. *Obes Rev.* 2015;16(8):621-638. doi:10.1111/obr.12288
5. Castaneda C, Marsden K, Maxwell T, et al. Prevalence of maternal obesity at delivery and association with maternal and neonatal outcomes. *J Matern Fetal Neonatal Med.* 2022;35(25):8544-8551. doi:10.1080/14767058.2021.1988563
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(1):5-14.
7. Xia Y, Xia C, Wu L, Li Z, Li H, Zhang J. Systemic Immune Inflammation Index (SII), System Inflammation Response Index (SIRI) and risk of all-cause mortality and cardiovascular mortality: a 20-year follow-up cohort study of 42,875 US adults. *J Clin Med.* 2023;12(3)doi:10.3390/jcm12031128
8. Zinellu A, Collu C, Nasser M, et al. The Aggregate Index of Systemic Inflammation (AIS): a novel prognostic biomarker in idiopathic pulmonary fibrosis. *J Clin Med.* 2021;10(18)doi:10.3390/jcm10184134
9. Ozkavak OO, Tanacan A, Serbetci H, et al. Association of first-trimester combined inflammatory markers (NLR, SII, SIRI, and AISI) with poor obstetric outcomes in pregnancies with ulcerative colitis. *Am J Reprod Immunol.* 2025;93(2):e70055. doi:10.1111/aji.70055

10. Marseglia L, Manti S, D'Angelo G, et al. Oxidative stress in obesity: a critical component in human diseases. *Int J Mol Sci.* 2015;16(1):378-400.
11. Zhang CXW, Candia AA, Sferuzzi-Perri AN. Placental inflammation, oxidative stress, and fetal outcomes in maternal obesity. *Trends Endocrinol Metab.* 2024;35(7):638-647. doi:10.1016/j.tem.2024.02.002
12. Shook LL, Kislal S, Edlow AG. Fetal brain and placental programming in maternal obesity: a review of human and animal model studies. *Prenat Diagn.* 2020;40(9):1126-1137. doi:10.1002/pd.5724
13. Vaughan OR, Rosario FJ, Chan J, et al. Maternal obesity causes fetal cardiac hypertrophy and alters adult offspring myocardial metabolism in mice. *J Physiol.* 2022;600(13):3169-3191. doi:10.1113/jp282462
14. Loh HH, Taipin H, Said A. Maternal obesity and risk of adverse obstetric outcomes in Malaysia. *Ann Acad Med Singap.* 2021;50(3):285-288. doi:10.47102/annals-acadmedsg.202054
15. Cardona-Benavides I, Mora-González P, Pineda A, Puertas A, Manzanares Galán S. Maternal obesity and the risk of fetal acidosis at birth. *J Matern Fetal Neonatal Med.* 2022;35(4):765-769. doi:10.1080/14767058.2020.1731795
16. Liu L, Ma Y, Wang N, Lin W, Liu Y, Wen D. Maternal body-mass index and risk of neonatal adverse outcomes in China: a systematic review and meta-analysis. *BMC Pregnancy Childbirth.* 2019;19(1):105. doi:10.1186/s12884-019-2249-z
17. Zhang Y, Lu M, Yi Y, et al. Influence of maternal body mass index on pregnancy complications and outcomes: a systematic review and meta-analysis. *Front Endocrinol (Lausanne.)* 2024;15:1280692. doi:10.3389/fendo.2024.1280692
18. Zhou Y, Wang Y, Wu T, Zhang A, Li Y. Association between obesity and systemic immune inflammation index, systemic inflammation response index among US adults: a population-based analysis. *Lipids Health Dis.* 2024;23(1):245. doi:10.1186/s12944-024-02240-8
19. Liao Y, Zhou K, Lin B, Deng S, Weng B, Pan L. Associations between Systemic Immune-Inflammatory Index and visceral adipose tissue area: results of a national survey. *Front Nutr.* 2024;11:1517186. doi:10.3389/fnut.2024.1517186
20. Gayret OB, Erol M, Buke O, Ozel A, Tenekeçigil A, Altuntaş MJJP. Systemic Inflammation Response Index and Immune-Inflammation Index as predictors of metabolic syndrome in obese children and adolescents: a prospective cross-sectional study. *Iran J Pediatr.* 2025;35(2):e148712.
21. Yildiz EG, Tanacan A, Okutucu G, Bastemur AG, Ipek G, Sahin D. Can System Inflammation Response Index or Systemic Immune Inflammation Index predict gestational diabetes mellitus in the first trimester? A prospective observational study. *Int J Gynaecol Obstet* 2024;166(2):837-843. doi:10.1002/ijgo.15442
22. Karatas E, Tanacan A, Ozkavak OO, et al. Predictive value of first-trimester Aggregate Index of Systemic Inflammation (AISI) and other inflammatory indices for gestational diabetes mellitus and associated obstetric outcomes. *Am J Reprod Immunol.* 2025;93(4):e70069. doi:10.1111/aji.70069
23. Lin Y, Fang J, Ni R, et al. Dynamic change of novel systemic inflammation markers to predict maternal-neonatal prognosis after cervical cerclage. *J Inflamm Res.* 2023;16:1745-1756. doi:10.2147/jir.S410211
24. Thomas MR, Storey RF. The role of platelets in inflammation. *Thromb Haemost.* 2015;114(3):449-458. doi:10.1160/th14-12-1067
25. Franco AT, Corken A, Ware J. Platelets at the interface of thrombosis, inflammation, and cancer. *Blood.* 2015;126(5):582-588. doi:10.1182/blood-2014-08-531582
26. Furuncuoğlu Y, Tulgar S, Dogan AN, Cakar S, Tulgar YK, Cakiroglu B. How obesity affects the neutrophil/lymphocyte and platelet/lymphocyte ratio, Systemic Immune-Inflammatory Index and platelet indices: a retrospective study. *Eur Rev Med Pharmacol Sci.* 2016;20(7):1300-1306.
27. Derhem B, Karahan İ. Early inflammation related to pediatric obesity. *J Pediatr Inf* 2023;17(1):9-13. doi:10.5578/ced.20239902
28. Luo L, Chen L, Song J, Ma X, Wang X. Association between Systemic Immune-Inflammatory Index and Systemic Inflammatory Response Index with body-mass index in children and adolescents: a population-based study based on the National Health and Nutrition Examination Survey 2017-2020. *Front Endocrinol (Lausanne).* 2024;15:1426404. doi:10.3389/fendo.2024.1426404
29. Sahin R, Tanacan A, Serbetci H, et al. First-trimester neutrophil-to-lymphocyte ratio (NLR), Systemic Immune-Inflammation Index (SII), and Systemic Immune-Response Index (SIRI) as predictors of composite adverse outcomes in pregnant women with Familial Mediterranean fever. *Z Geburthilfe Neonatol.* 2024;228(2):156-160. doi:10.1055/a-2125-0973
30. Sahin R, Tanacan A, Serbetci H, et al. The role of first-trimester NLR (neutrophil to lymphocyte ratio), Systemic Immune-Inflammation Index (SII), and Systemic Immune-Response Index (SIRI) in the prediction of composite adverse outcomes in pregnant women with systemic lupus erythematosus. *J Reprod Immunol.* Aug 2023;158:103978. doi:10.1016/j.jri.2023.103978
31. Zheng L, Ge R, Weng X, Lin L. Predictive value of serum immune-inflammatory markers for adverse pregnancy outcomes in pregnant women with thrombophilia: a retrospective cohort study. *J Inflamm Res.* 2024;17:6083-6091. doi:10.2147/jir.S481508
32. Sahin R, Tanacan A, Serbetci H, et al. The Association of Systemic Immune-Inflammation Index (SII), Systemic Immune-Response Index (SIRI), and neutrophil-to-lymphocyte ratio (NLR) with cesarean scar pregnancy (CSP). *J Reprod Immunol.* 2024;164:104275. doi:10.1016/j.jri.2024.104275

Visual recovery after phacoemulsification surgery: assessing the role of age and gender

 Ceren Türkoğlu

Department of Ophthalmology, Faculty of Medicine, Lokman Hekim University, Ankara, Türkiye

Cite this article as: Türkoğlu C. Visual recovery after phacoemulsification surgery: assessing the role of age and gender. *J Health Sci Med.* 2025;8(6):1129-1132.

Received: 05.09.2025

Accepted: 12.10.2025

Published: 25.10.2025

ABSTRACT

Aims: There is a need for more comprehensive resources on whether visual outcomes after phacoemulsification surgery are affected by age and gender. This study investigates visual changes in relation to age and gender following this procedure. This study aims to explore the differences in visual recovery after phacoemulsification surgery between patients aged 50-69 and those aged 70 and above, with a focus on gender-related differences in outcomes.

Methods: This retrospective, single-center, cross-sectional study included 185 eyes of 185 patients aged 50 and above who underwent phacoemulsification. Patients were divided into two groups; group 1 (50-69 years) and group 2 (70+ years). Preoperative best-corrected visual acuity (BCVA), intraocular pressure, and other ocular parameters were measured. Postoperative follow-up occurred at 1, 3, 6, and 12 months. Statistical analyses were performed to evaluate differences in visual recovery between age groups and genders.

Results: Statistically significant improvements in BCVA and IOP were observed between preoperative and all postoperative time points ($p < 0.001$). However, no significant differences were found in BCVA improvement, IOP changes, or complication rates between the two age groups or between male and female patients. The complication rate was 3.8% intraoperatively and 1.6% postoperatively, with no significant gender-related differences. These findings suggest that with a standardized surgical approach, patient outcomes are consistent. This highlights the importance of surgical experience in patient counseling and surgical planning, indicating that excellent results can be expected regardless of age or gender.

Conclusion: Phacoemulsification surgery leads to significant improvements in visual acuity and IOP. The experience of the surgeon may play a pivotal role in minimizing complications and ensuring similar positive outcomes for both male and female patients across various age groups. Larger, multicenter studies are recommended to further validate these findings and explore outcomes in patients with comorbidities.

Keywords: Phacoemulsification, visual acuity, gender, age, intraocular pressure

INTRODUCTION

Senile cataract is a leading cause of reversible blindness worldwide, predominantly affecting individuals over 50. The condition arises from the opacification of the lens due to age-related changes in lens proteins.¹ While typically bilateral, its progression can be asymmetric. Key risk factors include age, female gender, ultraviolet light exposure, diabetes mellitus, smoking, corticosteroid use, and a history of ocular trauma or surgery.^{1,2} As patients age, factors such as harder cataracts, shallower anterior chambers, smaller pupils, higher rates of exfoliation syndrome, and weaker zonules can make surgery more challenging.²

Phacoemulsification has become the standard surgical method for cataracts, favored for being less traumatic, having a shorter surgical time, and causing fewer complications than older techniques. This modern procedure involves removing the opacified lens through ultrasonic fragmentation and

implanting a clear intraocular lens, all within a closed system maintained by small incisions.³

The visual system naturally undergoes age-related functional decline. For instance, retinal ganglion cell axons are vulnerable to age-related loss, and the total axon number in the optic nerve is inversely proportional to aging.⁴ However, the specific impact of age and gender on visual recovery after phacoemulsification remains an area with insufficient research. This study aims to address this gap by investigating vision changes in two age groups (50-69 and ≥ 70 years) and between genders. The primary purpose is to determine if age and gender are significant predictors of visual recovery, thereby providing valuable data for patient counseling and managing postoperative expectations. The novelty of this study lies in its simultaneous investigation of both age and gender, offering a more integrated understanding of postoperative outcomes.

Corresponding Author: Ceren Türkoğlu, dr.cerenturkoglu@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

METHODS

This retrospective, cross-sectional study was approved by the Lokman Hekim University Scientific Researches Ethics Committee (Date: 30.05.2025, Decision No: 2025/153) and adhered to the principles of the Declaration of Helsinki. The records of 185 eyes from 185 patients aged 50 and over who underwent phacoemulsification at our ophthalmology clinic over a 12-month period were analyzed. Patients with nuclear, cortical, and/or posterior subcapsular cataracts (grade 2-4, LOCS III) were included. Exclusion criteria were the presence of other ocular diseases (e.g., glaucoma, diabetic retinopathy), a history of ocular trauma, or previous ocular surgery.

Patients were divided into two age groups; group 1 (n=93, 50-69 years) and group 2 (n=92, ≥70 years). All patients underwent a comprehensive preoperative assessment, including best-corrected visual acuity (BCVA) measurement using LogMAR charts, slit-lamp biomicroscopy, detailed fundus examination, and optical biometry. Spectral domain-optical coherence tomography (SD-OCT) was used to exclude macular pathologies. A single experienced surgeon performed all phacoemulsification and monocular intraocular lens implantations to minimize inter-surgeon variability. Postoperative follow-ups were conducted at 1, 3, 6, and 12 months, recording the same parameters.

Statistical Analysis

All data analysis was performed using SPSS version 23 (SPSS Inc., Chicago, Illinois, USA). The normality of data distribution was assessed using the Shapiro-Wilk test. For normally distributed data, comparisons between two independent groups were made using the student's T test. For non-normally distributed data, the Mann-Whitney U test was applied. Paired T tests were used to compare preoperative and postoperative values within the same group. A p-value of <0.05 was considered statistically significant. Based on a post hoc power analysis with α=0.05, power=0.80, and a medium effect size, the minimum required sample size was 88 eyes per group. The study included 185 eyes in total, meeting this requirement.

RESULTS

Out of the 185 patients, 185 eyes were included. The mean age of the patients was 68.30±0.51 and 94 (50.8%) were female and 91 (49.2%) were male. Ninety (48.6%) of the patients underwent surgery on the right eye and 95 (51.4%) on the left eye. The mean age of the 93 patients in group 1 was 62.81±0.47 and 51 (54.8%) were female and 42 (45.2%) were male. Of the patients in group 1, 48 (51.6%) were operated on the right eye

and 45 (48.4%) on the left eye. The mean age of the 92 patients in group 2 was 73.84±0.41, and 43 (46.7%) of the patients were female and 49 (53.3%) were male. Of the patients in group 2, 42 patients (45.7%) underwent surgery on the right eye and 50 patients (54.3%) on the left eye.

Considering the pre-operative (pre-op) refractive properties of the patients, the spherical value of 185 patients was -1.32±2.48, the cylindrical value was -1.19±0.80, and the cylindrical axis value was 83.83±51.34. The spherical value of the patients in group 1 was -1.32±2.65, the cylindrical value was -1.13±0.81, and the cylindrical axis value was 81.69±55.66. The spherical value of the patients in group 2 was -1.32±2.32, the cylindrical value was -1.26 ±0.79, and the cylindrical axis value was 85.98±46.77. The postoperative (post-op) 12th month spherical value was determined as 0.11±0.31 and there was a statistically significant difference with the pre-op value (p<0.001). The post-op cylindrical value was determined as -0.70±0.04 and there was a statistically significant difference with the pre-op value (p<0.001). The cylindrical axis value was determined as 82.68±3.99 and there was no statistically significant difference with the pre-op value (p=0.8).

Considering BCVA changes for all patients during phacoemulsification surgery, a statistically significant difference was found between preoperative BCVA and post-operative 1-month, post-operative 3-month, post-operative 6-month and post-operative 12-month periods (p<0.001). Similar statistically significant differences were observed in both group 1 and group 2 patients between preoperative BCVA and all postoperative time points (p<0.001). The BCVA changes of the patients over 12 months are given in [Table 1](#).

No statistically significant difference was found between pre-operative and post-operative 1, 3, 6 and 12-month BCVA changes of the patients (p=0.47, 0.64, 0.66, 0.67, respectively). In addition, when groups 1 and 2 were compared, it was found that there was no statistically significant between pre-operative and post-operative 1, 3, 6 and 12-month BCVA changes (p=0.34, 0.26, 0.24, 0.25, respectively).

Pre-postoperative intraocular pressure values of the patients were compared. A statistically significant difference was found between preoperative IOP and postoperative 1-month, 3-month, 6-month, and 12-month IOP values (p<0.001). Examination of both group 1 and group 2 revealed a statistically significant difference between preoperative IOP and postoperative 1-month, 3-month, 6-month, and 12-month IOP values (p<0.001). The IOP changes of the patients are given in [Table 2](#).

Table 1. Changes of BCVA values throughout 12 months

	Pre-op	Post-op 1-month	Post-op 3-month	Post-op 6-month	Post-op 12-month
All patients (n=185)	0.66±0.36	0.12±0.23 p<0.001	0.05±0.06 p<0.001	0.04±0.06 p<0.001	0.04±0.06 p<0.001
Group 1 (n=93)	0.62±0.34	0.09±0.09 p<0.001	0.03±0.05 p<0.001	0.03±0.05 p<0.001	0.03±0.05 p<0.001
Group 2 (n=92)	0.71±0.36	0.15±0.31 p<0.001	0.06±0.07 p<0.001	0.06±0.07 p<0.001	0.06±0.07 p<0.001

Pre-op: Pre-operative, Post-op: Post-operative, BCVA: Best corrected visual acuity

Table 2. Changes of IOP values throughout 12 months

IOP values	Pre-op	Post-op 1-month	Post-op 3-month	Post-op 6-month	Post-op 12-month
All patients	15.22±2.51	14.08±2.45 p<0.001	13.66±2.18 p<0.001	13.23±2.16 p<0.001	12.92±2.05 p<0.001
Group 1	14.98±2.52 p<0.001	13.95±2.54 p<0.001	13.49±2.29 p<0.001	13.12±2.31 p<0.001	12.75±2.19 p<0.001
Group 2	15.46±2.48 p<0.001	14.22±2.37 p<0.001	13.84±2.06 p<0.001	13.35±2.01 p<0.001	13.10±1.90 p<0.001

IOP: Intraocular pressure, Pre-op: Pre-operative, Post-op: Postoperative

Considering the pre-operative and post-operative IOP values no statistically significant difference was found between male and female patients ($p>0.05$). In female patients, a statistically significant difference was found between pre-operative IOP and post-operative 1-month, 3-month, 6-month and 12-month IOP values ($p<0.001$). Similarly, in male patients, a statistically significant difference was found between preoperative IOP and postoperative 1-month, 3-month, 6-month, and 12-month IOP values ($p<0.001$).

Regarding the complication rate, 3.8% of patients developed intraoperative complications (posterior capsule opening, zonule dialysis). While 1.6% of intra-operative complications occurred in group 1, 2.2% in group 2. Postoperative complications (seidel) were observed in 1.6% of the patients. While 1.1% of post-operative complications occurred in group 1, 0.5% in group 2.

DISCUSSION

This study investigated the impact of age and gender on visual recovery after phacoemulsification surgery and found that neither factor significantly influenced postoperative outcomes in our cohort. Significant improvements in both BCVA and IOP were observed across all patients, reinforcing the high efficacy and safety of modern cataract surgery. The key finding is that when surgery is performed by an experienced surgeon on patients without significant ocular comorbidities, excellent and comparable visual recovery can be expected across different age groups and genders.

Our results align with some previous studies. For instance, Ben-Eli et al.⁵ also reported similar visual outcomes for younger and older patients after phacoemulsification. While some literature suggests older patients may face higher risks due to factors like zonular weakness or smaller pupils,^{6,7} our findings indicate that these age-related challenges may not translate to poorer outcomes when managed by an experienced surgeon. Similarly, a study by Toyama et al.² found comparable visual acuity and complication rates between patients <80 and ≥90 years old, though they noted less favorable outcomes in patients with diabetes. Our study, by excluding such comorbidities, isolates the effect of age and demonstrates its limited impact on recovery under ideal conditions.

The role of gender in cataract surgery outcomes remains debated. Some studies have suggested women may experience worse outcomes, potentially due to anatomical or hormonal differences.^{8,9} However, our findings are consistent with other research, such as studies by Ben-Eli et al.⁵ and Lundqvist et

al.,¹⁰ which found no significant gender-based differences in visual acuity improvement. The standardized technique used by a single surgeon in our study likely minimized variability, contributing to the uniform outcomes between genders. Nevertheless, certain baseline ocular characteristics-such as smaller average pupil size and greater nuclear sclerosis reported in women-could theoretically influence surgical difficulty and visual recovery. Although these parameters were not analyzed in our cohort, they may partly explain gender differences observed in other studies and warrant investigation in future research.

A significant reduction in IOP was a consistent finding, aligning with previous literature that attributes this effect to improved aqueous humor outflow after lens removal.^{11,12} The low complication rates (3.8% intraoperative, 1.6% postoperative) are comparable to those reported in large-scale studies and further underscore the procedure's safety.^{5,13} The absence of significant differences in complications between age and gender groups reinforces that phacoemulsification is a robust procedure for a diverse patient population.

The clinical implication of these findings is significant for patient counseling. Surgeons can confidently inform patients, regardless of their age (within the studied range) or gender, that they have a high probability of achieving excellent visual outcomes, provided there are no other underlying ocular diseases. This helps manage patient expectations and reinforces the reliability of phacoemulsification.

Limitations

This study has several limitations that must be acknowledged. First, its retrospective, single-center design inherently limits the generalizability of the findings. The results may not be applicable to settings with different patient demographics or surgical expertise. Second, the exclusion of patients with significant ocular or systemic comorbidities, while reducing confounding variables, also limits the external applicability of our findings to the general cataract population, which often presents with such conditions. Other limitations include the lack of analysis on gender-related baseline parameters like nuclear hardness or pupil size, and a follow-up period limited to 12 months; longer-term outcomes remain unassessed. These parameters could be relevant to gender differences, for example, smaller pupil size in women or variations in nuclear sclerosis. Future research should investigate these factors to provide a more comprehensive understanding of gender-specific outcomes. These limitations should be addressed in future research.

CONCLUSION

In this study, age and gender did not significantly affect postoperative visual acuity improvement, IOP changes, or complication rates following phacoemulsification surgery. Both younger (50-69 years) and older (≥ 70 years) patients, as well as male and female patients, demonstrated substantial and comparable improvements in visual outcomes. These findings suggest that phacoemulsification, when performed by an experienced surgeon, is a highly effective and safe procedure for a broad spectrum of cataract patients. Future research should include larger, multicenter prospective studies with more diverse patient cohorts, including those with comorbidities, to validate these results and provide a more comprehensive understanding of postoperative recovery.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Lokman Hekim University Scientific Researches Ethics Committee (Date: 30.05.2025, Decision No: 2025/153).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Saad Filho R, Moreto R, Nakaghi RO, Haddad W, Coelho RP, Messias A. Costs and outcomes of phacoemulsification for cataracts performed by residents. *Arq Bras Oftalmol.* 2020;83(3):209-214. doi:10.5935/0004-2749.20200059
2. Toyama T, Ueta T, Yoshitani M, Sakata R, Numaga J. Visual acuity improvement after phacoemulsification cataract surgery in patients aged ≥ 90 years. *BMC Ophthalmol.* 2018;18(1):280. doi:10.1186/s12886-018-0950-8
3. Lu Q, Wei L, He W, et al. Lens Opacities Classification System III-based artificial intelligence program for automatic cataract grading. *J Cataract Refract Surg.* 2022;48(5):528-534. doi:10.1097/j.jcrs.0000000000000790
4. Calkins DJ. Age-related changes in the visual pathways: blame it on the axon. *Invest Ophthalmol Vis Sci.* 2013;54(14):ORSF37-ORSF41. doi:10.1167/iops.13-12784
5. Ben-Eli H, Cnaany Y, Halpert M, Chowers I, Goldstein A. Investigating the impact of age and sex on cataract surgery complications and outcomes. *Sci Rep.* 2025;15(1):1242. doi:10.1038/s41598-024-84382-4
6. Mönestam E, Wachmeister L. Impact of cataract surgery on the visual ability of the very old. *Am J Ophthalmol.* 2004;137(1):145-155. doi:10.1016/s0002-9394(03)00900-0
7. Marmamula S, Barrenakala NR, Challa R, et al. Visual outcomes after cataract surgery among the elderly residents in the 'homes for the aged' in South India: the Hyderabad ocular morbidity in elderly study. *Br J Ophthalmol.* 2021;105(8):1087-1093. doi:10.1136/bjophthalmol-2020-317167
8. Hashemi H, Mohammadi SF, Z-Mehrjardi H, et al. The role of demographic characteristics in the outcomes of cataract surgery and gender roles in the uptake of postoperative eye care: a hospital-based study. *Ophthalmic Epidemiol.* 2012;19(4):242-248. doi:10.3109/09286586.2012.691600
9. Quintana JM, Garcia S, Aguirre U, et al. Relationship of sociodemographic variables with outcomes after cataract surgery. *Eye (Lond).* 2013;27(6):698-708. doi:10.1038/eye.2013.85
10. Lundqvist B, Mönestam E. Gender-related differences in cataract surgery outcome: a 5-year follow-up. *Acta Ophthalmol.* 2008;86(5):543-548. doi:10.1111/j.1600-0420.2007.01099.x
11. Hayashi K, Hayashi H, Nakao F, Hayashi F. Effect of cataract surgery on intraocular pressure control in glaucoma patients. *J Cataract Refract Surg.* 2001;27(11):1779-1786. doi:10.1016/s0886-3350(01)01036-7
12. Rothman AL, Chang TC, Lum F, Vanner EA. Intraocular pressure changes following stand-alone phacoemulsification: an IRIS-registry analysis. *Am J Ophthalmol.* 2023;245:25-36. doi:10.1016/j.ajo.2022.09.006
13. Kim BZ, Patel DV, McGhee CN. Auckland cataract study 2: clinical outcomes of phacoemulsification cataract surgery in a public teaching hospital. *Clin Exp Ophthalmol.* 2017;45(6):584-591. doi:10.1111/ceo.12922

Retrospective evaluation of the demographic, psychiatric, and clinical characteristics of patients admitted to the intensive care unit following a suicide attempt

 Meltem Hazel Şimşek,  Ulaş Korkmaz

Department of Psychiatry, Faculty of Medicine, Giresun University, Giresun, Türkiye

Cite this article as: Şimşek MH, Korkmaz U. Retrospective evaluation of the demographic, psychiatric, and clinical characteristics of patients admitted to the intensive care unit following a suicide attempt. *J Health Sci Med.* 2025;8(6):1133-1142.

Received: 28.07.2025

Accepted: 13.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This retrospective, multicenter study aims to evaluate the relationship between demographic, clinical, and psychiatric characteristics and mortality among patients admitted to the intensive care unit (ICU) following a suicide attempt.

Methods: Between January 2020 and May 2025, a total of 121 patients admitted to the ICU following a suicide attempt at three hospitals in a province of Türkiye were retrospectively reviewed. Demographic, clinical, and psychiatric data were extracted from medical records and analyzed to assess their associations with mortality.

Results: Male sex (with 96.3% of deceased cases being male) (aOR=35.402, 95% CI=1.821-688.370), jumping from a height (aOR=57.931, 95% CI=2.912-1152.486), and the absence of psychiatric consultation during ICU stay (aOR=150.134, 95% CI=10.590-2128.469) were all significantly associated with mortality. Patients who used the lethal methods had more extended ICU stays and were more frequently referred to locked psychiatric units. Mortality rates were significantly higher among those who did not receive a psychiatric consultation.

Conclusion: This retrospective study investigated key risk factors associated with mortality among patients admitted to the ICU following a suicide attempt. Advanced age, male sex, the presence of physical illness, use of high-lethality suicide methods, and the absence of psychiatric consultation were found to be the major factors that may contribute to mortality in this population. The findings suggest the critical importance of early identification of high-risk individuals, timely psychiatric assessment, and a multidisciplinary approach to intervention in reducing suicide-related mortality, although the results should be interpreted with caution given the limited number of events and retrospective design. Additionally, potential time-to-consultation biases warrant cautious interpretation.

Keywords: Suicide attempt, intensive care, mortality, psychiatric consultation

INTRODUCTION

Suicide is defined as the act of deliberately and voluntarily engaging in behavior that may result in one's death.¹ It represents a growing public health concern with significant consequences for individuals and societies alike.² According to the latest data from the World Health Organization, approximately 800,000 to 1 million people die by suicide each year.³ Suicide is the second most common cause of death among individuals aged 15 to 29.⁴

Suicidal behavior encompasses a continuum that includes suicidal ideation, suicide attempts, and completed suicide.⁵ While suicide attempts are more frequently reported among women, completed suicides occur approximately three times more often in men.⁶ Although suicide attempts are more common among young people and women, their prevalence is increasingly rising among older adults and men. In older populations, suicide attempts are more likely to result in death compared to those in younger individuals.⁷

Studies have shown that the most common methods of suicide attempts include drug or toxic substance ingestion, hanging, jumping from a height, and the use of firearms or sharp objects.⁸ In the general population and across both sexes, drug intoxication is the most frequently reported method. However, suicide attempts involving hanging, jumping from a height, firearms, and sharp instruments are reported more frequently among men and older adults compared to women and younger individuals.^{8,9}

Some suicide attempts are particularly severe and life-threatening, requiring monitoring and treatment in the intensive care unit. A significant proportion of high-lethality attempts are accompanied by underlying psychiatric disorders.¹⁰⁻¹² Moreover, many of these high-lethality cases involve comorbidity with multiple psychiatric conditions.¹³

One of the key risk factors for serious suicide attempts is a prior history of suicide attempts. A previous suicide attempt

Corresponding Author: Ulaş Korkmaz, ulas.korkmaz@giresun.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

is among the strongest predictors of future attempts and death by suicide.¹⁴ Another factor influencing suicide risk is the individual's social support system. Literature indicates that suicide rates are significantly lower among married individuals compared to those who are unmarried. Conversely, individuals who are divorced or widowed have been found to have a markedly higher risk of suicide.¹⁵ Additionally, the presence of physical illness is another important risk factor. Studies have shown that individuals with physical illnesses are more likely to engage in severe and fatal suicide attempts compared to those without such conditions.^{16,17}

Suicide attempts may vary depending on an individual's place of residence. Rates of suicide have been found to be higher in rural areas compared to urban settings.¹⁸ Furthermore, individuals living in rural regions are more likely to employ more lethal methods of suicide.¹⁹ This pattern has been linked to reduced access to psychiatric services and greater difficulty obtaining appropriate treatment in rural communities^{18,19}. The frequency and lethality of suicide attempts also appear to fluctuate seasonally. Studies have shown that suicide attempts increase during the spring and summer months, with a decline observed in the autumn.^{20,21} Attempt rates tend to peak in the summer among women and younger individuals, while men are more likely to engage in violent and severe attempts during the spring²². Cross-national studies have also indicated that completed suicides occur most frequently in spring and follow a recurring seasonal pattern.^{20,23,24}

As the severity of a suicide attempt increases, so does the need for intensive care and the length of the intensive care unit (ICU) stay.²⁵ High-lethality methods such as jumping from a height, hanging, severe multi-trauma, and ingestion of multiple or highly toxic substances often necessitate admission to the ICU. The duration of ICU stay is typically prolonged depending on the severity of the attempt, the extent of resulting organ failure, and the presence of additional medical complications.^{26,27}

Studies have reported that some patients admitted to the ICU due to suicide attempts do not receive a psychiatric evaluation during their hospitalization.^{28,29} However, the absence or delay of psychiatric assessment may increase the likelihood of repeated suicide attempts and hinder access to appropriate treatment.^{30,31} Therefore, a timely psychiatric evaluation plays a critical role in the post-attempt treatment process.

The literature suggests that, following medical stabilization after a suicide attempt, assigning patients to open or locked psychiatric units based on psychiatric assessment outcomes can help reduce the likelihood of repeated attempts.^{32,33} However, some individuals may refuse psychiatric evaluation due to lack of insight, poor treatment adherence, or fear of stigma, and consequently remain outside the scope of follow-up care.³⁴ On the other hand, for those without active suicidal ideation, severe psychiatric symptoms, and who demonstrate good treatment adherence and social support, outpatient follow-up may be considered sufficient.³⁵ There is limited data in the literature regarding the psychiatric assessment processes of patients admitted to the ICU following a suicide attempt, including referral rates to psychiatry, whether patients were admitted to open or locked units, and rates of treatment

refusal. Moreover, studies that comprehensively examine these patients' sociodemographic, clinical, and attempt-related characteristics alongside psychiatric consultation status and clinical outcomes (e.g., death, psychiatric hospitalization, or outpatient referral) remain scarce.

This study retrospectively analyzed the demographic, clinical, and psychiatric characteristics of patients admitted to the ICU following a suicide attempt between January 2020 and May 2025. It aims to evaluate the associations between these characteristics and whether patients received psychiatric consultation, were referred to open or locked psychiatric units, and their clinical outcomes (e.g., death, discharge, or inpatient admission). Additionally, the study aims to identify variables that are mainly associated with mortality within this patient population.

The hypotheses tested in this study regarding patients admitted to the ICU following a suicide attempt are as follows:

- **H1:** Demographic characteristics are associated with clinical outcomes.
- **H2:** Psychiatric history (including previous diagnoses, psychotropic medication use, and past suicide attempts) influences clinical outcomes.
- **H3:** The presence of physical illness is associated with higher mortality rates following a suicide attempt.
- **H4:** The method of suicide attempt is associated with mortality.
- **H5:** There are differences in mortality rates and discharge types between patients who received psychiatric consultation and those who did not.
- **H6:** Length of ICU stay is associated with clinical factors such as the method of suicide attempt and comorbid physical illness.

METHODS

Ethics

The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Ethical approval was obtained from the Giresun Training and Research Hospital Ethics Committee (Date: 23.07.2025, Decision No: 23.07.2025/14). Ethical approval was obtained from this single center only, and permission to access and use data from the other participating institutions was obtained from the Provincial Directorate of Health. Given the retrospective design, the requirement for individual informed consent was waived by the ethics committee.

Study Design

This study is a retrospective, descriptive, and comparative analysis of patients who were admitted to the ICU due to suicide attempts at three hospitals in a province of Türkiye between January 2020 and May 2025.

Medical records of all eligible patients were reviewed retrospectively. Sociodemographic (e.g., age, sex, marital status), clinical (e.g., length of ICU stay, presence of comorbid physical illness-chronic diseases such as diabetes,

hypertension-, method of suicide attempt), and psychiatric data (e.g., known psychiatric diagnoses, psychotropic medication use, history of previous suicide attempts) were extracted from patient files. Cases with missing data, those determined not to involve a suicide attempt, or those lacking sufficient clinical documentation to confirm the presence or absence of a suicide attempt were excluded from the study.

Inclusion criteria were as follows; being 18 years of age or older, admission to the ICU due to a suicide attempt, clear documentation of the suicide attempt in the clinical records, and availability of data regarding the ICU course, treatment, psychiatric evaluation, and clinical outcome. Exclusion criteria included: patients admitted to the ICU for reasons other than suicide attempts, cases with significant missing data in the medical records (e.g., absence of diagnostic information, method of attempt, or outcome), patients for whom the presence of a suicide attempt could not be clinically confirmed or remained uncertain, and individuals under the age of 18.

Statistical Analysis

Data analysis was conducted using IBM SPSS Statistics version 27. The normality of continuous variables was assessed based on skewness and kurtosis values, with values between ± 1.5 considered indicative of normal distribution. Descriptive statistics were presented as mean $\pm$ standard deviation or median (interquartile range, IQR) for continuous variables, and as frequency (n) and percentage (%) for categorical variables.

Group comparisons were performed using independent samples T tests when parametric assumptions were met, and Mann-Whitney U or Kruskal-Wallis H tests when they were not. For categorical variables, group comparisons were conducted using the chi-square test or Fisher's exact test, as appropriate. Bonferroni correction was made for multiple comparisons. Additionally, relative risk, risk differences, and effect sizes were reported.

To identify variables associated with mortality following suicide attempts, logistic regression analyses were performed. In these analyses, mortality was the dependent variable, and independent variables included age, sex, marital status, place of residence, history of psychiatric diagnosis and treatment, presence of physical illness, method of suicide attempt, length of ICU stay, history of previous suicide attempts, and whether psychiatric consultation was conducted. Odds ratios and corresponding 95% confidence intervals were reported for each variable.

A p-value of <0.05 was considered statistically significant in all analyses.

RESULTS

A total of 137 patient records were identified through data screening. Following the application of inclusion and exclusion criteria, 121 patients were included in the final analysis (Figure).

Table 1 compares the sociodemographic characteristics of patients admitted to the ICU following a suicide attempt,

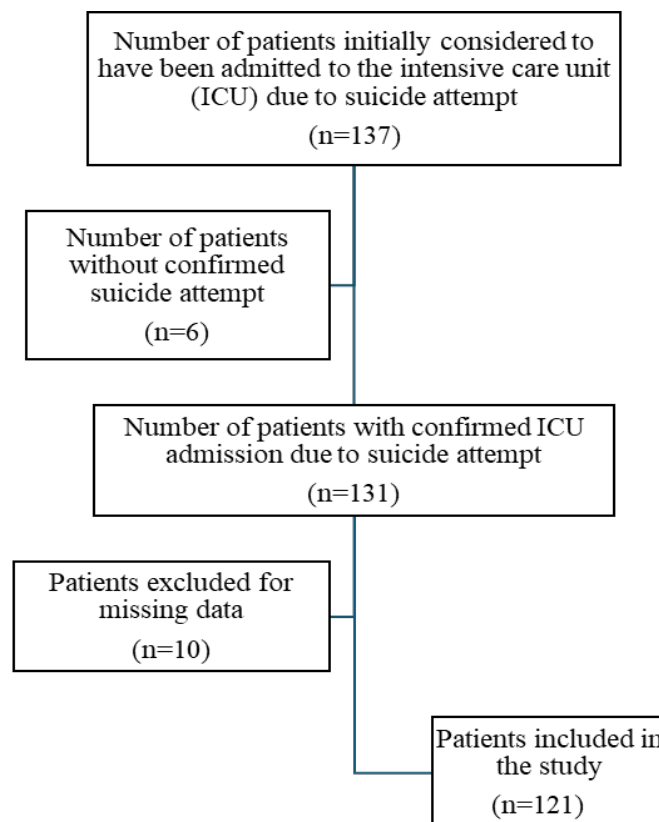


Figure. Study flow diagram

stratified by survival status. The mean age of deceased patients was significantly higher than that of survivors ($p=0.036$). There was also a significant difference in gender distribution: 96.3% of the deceased group were male, compared to approximately 48% among survivors ($p<0.001$). In contrast, no significant differences were found between the groups in terms of marital status, place of residence, or the season and year in which the attempt occurred (all $p>0.05$). Nonetheless, it is notable that the majority of suicide attempts occurred between 2021 and 2023 and were most frequent in the spring season.

Table 2 compares the clinical characteristics of survivors and deceased patients following admission to the ICU after a suicide attempt. No significant difference was found between the two groups in terms of ICU length of stay ($p=0.132$). Similarly, there were no statistically significant differences in the presence of a prior psychiatric diagnosis ($p=1.00$), psychotropic medication use ($p=0.847$), or history of previous suicide attempts ($p=1.00$). In contrast, the presence of a known physical illness was significantly associated with mortality; comorbid physical conditions were more common among deceased patients than among survivors ($p=0.019$). The method of suicide attempt also had a significant impact on clinical outcomes ($p=0.006$). While the majority of survivors had attempted suicide via drug ingestion (78.7%), this method was less common among deceased patients (37%). Conversely, methods such as hanging, jumping from a height, or the use of firearms/sharp objects were more prevalent among those who died. Moreover, there was a marked difference between groups regarding whether psychiatric consultation was requested. While 80.9% of survivors received a psychiatric consultation, only 11.1% of the deceased did ($p<0.001$).

Table 1. Comparison of sociodemographic characteristics between survivors and deceased patients (n=121)

	Survivors (n=94)	Deceased (n=27)	Statistic	p	Effect size (Cohen's d/Cramer's V)	Risk ratio (risk differences)
	Mean±SD/n (%)					
Age (years)	33.36±12.95	42.19±19.81	t=-2.184	0.036*		
Gender			X ² =18.337	<0.001**	0.389	18.3 (0.346)
Female	49 (52.1%)	1 (3.7%)				
Male	45 (47.9%)	26 (96.3%)				
Marital status			X ² =0.316	0.574**	0.051	1.36 (0.065)
Married	32 (34%)	7 (25.9%)				
Single	62 (66%)	20 (74.1%)				
Place of residence			X ² =1.478	0.224**	0.110	1.66 (0.109)
Urban	46 (48.9%)	9 (33.3%)				
Rural	48 (51.1)	18 (66.7%)				
Season of attempt			X ² =1.549	1***	0.113	
Autumn	14 (14.9%)	6 (22.2%)				
Winter	25 (26.6%)	5 (18.5%)				
Spring	30 (31.9%)	10 (37%)				
Summer	25 (26.6%)	6 (22.2%)				
Year of attempt			X ² =8.124	1****	0.259	
2020	7 (7.4%)	2 (7.4%)				
2021	22 (23.4%)	8 (29.6%)				
2022	20 (21.3%)	6 (22.2%)				
2023	26 (27.7%)	3 (11.1%)				
2024	9 (9.6%)	7 (25.9%)				
2025	10 (10.6%)	1 (3.7%)				

n: Number of cases, p: Significance level, t: T test statistic, X²: Chi-square statistic, *Student's T test, **Yates Chi-square test, ***Pearson Chi-square test, ****Fisher's exact test, SD: Standard deviation

Table 2. Comparison of clinical characteristics between survivors and deceased patients (n=121)

	Survivors (n=94)	Deceased (n=27)	Statistic	p	Effect size (r / Cramer's V)	Risk ratio (risk differences)
	Median (Q1-Q3)/n (%)					
ICU stay (days)	2 (1-4)	2 (1-9)	Z=-1.505	0.132*	-0.137	
Known psychiatric diagnosis			X ² <0.001	1**	0.003	1.07 (0.015)
No	26 (27.7)	7 (25.9%)				
Yes	68 (72.3%)	20 (74.1%)				
Psychotropic medication use			X ² =0.037	0.847**	0.018	0.87 (-0.032)
No	55 (58.5%)	17 (63%)				
Yes	39 (41.5%)	10 (37%)				
Comorbid physical illness			X ² =5.508	0.019**	0.213	2.39 (0.237)
No	78 (83%)	16 (59.3%)				
Yes	16 (17%)	11 (40.7%)				
Method of suicide attempt			X ² =16.125	0.006***	0.365	
Drug ingestion	74 (78.7%)	10 (37%)				
Hanging	5 (5.3%)	4 (14.8%)				
Jumping from a height	5 (5.3%)	4 (14.8%)				
Firearm/sharp object	10 (10.6%)	9 (33.3%)				
History of suicide attempt			X ² <0.001	1**	0	0.96 (-0.010)
No	58 (61.7%)	17 (63%)				
Yes	36 (38.3%)	10 (37%)				
Psychiatric consultation requested			X ² =41.991	<0.001**	0.589	0.067 (-0.533)
No	18 (19.1%)	24 (88.9%)				
Yes	76 (80.9%)	3 (11.1%)				

n: Number of cases, p: Significance level, Q1: First quartile, Q3: Third quartile, Z: Mann-Whitney U test statistic, X²: Chi-square statistic, ICU: Intensive care unit, r: Rank-biserial correlation, *Mann-Whitney U test, **Yates Chi-square test, ***Fisher's exact test

Table 3 presents the relationship between the type of discharge following ICU admission and both the method of suicide attempt and whether a psychiatric consultation was requested. A statistically significant association was found between discharge type and method of suicide attempt ($p<0.001$). For instance, all patients transferred to internal medicine wards had attempted suicide by drug ingestion. In contrast, more lethal methods, such as hanging, jumping from a height, or the use of firearms/sharp objects, were more frequently associated with death or referral to psychiatric units, particularly locked psychiatric units. Likewise, the vast majority of patients who were discharged in stable condition had attempted suicide by drug ingestion. Among those discharged against medical advice, the rate of suicide attempts involving sharp or penetrating objects was notably higher compared to other groups. There was also a significant association between discharge type and whether psychiatric consultation was requested ($p<0.001$). Nearly all patients referred to any psychiatric unit had received a psychiatric consultation during their ICU stay, while the majority of those who died had not. A substantial proportion of patients transferred to internal medicine or discharged in stable condition also received psychiatric consultations. In contrast, most of the patients discharged after refusing treatment had not been evaluated by psychiatry.

Table 4 examines the impact of suicide attempt method and the presence of comorbid physical illness on the length

of stay in the ICU. A statistically significant difference was not observed in ICU stay duration based on the method of suicide attempt ($p=0.102$). Nevertheless, patients who attempted suicide by hanging or jumping from a height had more extended ICU stays compared to those who used drug ingestion or sharp/firearm-related methods, which were associated with shorter median stays. The presence or absence of a known physical illness was not significantly associated with the duration of ICU stay ($p=0.393$).

Table 5 evaluates the independent variables predicting mortality following a suicide attempt using univariate logistic regression analysis. The results indicate that older age significantly associated with the likelihood of mortality (OR=1.04 per year increase, 95% CI=1.008-1.066, $p=0.011$). Similarly, a longer duration of ICU stay was also an independent risk factor for death (OR=1.11 per additional day, 95% CI=1.013-1.206, $p=0.025$). Sex emerged as a strong predictor of mortality: male patients were significantly more likely to die than female patients ($p=0.001$). In this model, being male was associated with an OR of 28.311 (95% CI=3.689-217.269), suggesting that the risk of death among men was approximately 28 times higher than among women. Furthermore, The presence of comorbid physical illness was significantly associated with increased mortality risk (OR=3.35, 95% CI=1.313-8.557, $p=0.011$). In contrast, variables such as marital status, place of residence, previous psychiatric diagnosis, and use of psychotropic medication

Table 3. Relationship between discharge type, suicide attempt method, and psychiatric consultation (n=121)

	Deceased (n=27)	Transferred to internal medicine (n=15)	Referred to psychiatric unit (n=17)	Referred to locked psychiatric unit (n=19)	Discharged in stable condition (n=31)	Discharged against medical advice (n=12)	Statistic	p	Effect size (Cramer's V)
Method of suicide attempt							$X^2=43.285$	$<0.001^*$	0.345
Drug ingestion	10 (37%)	15 (100%)	13 (76.5%)	10 (52.6%)	28 (90.3%)	8 (66.7%)			
Hanging	4 (14.8%)	0 (0%)	1 (5.9%)	3 (15.8%)	1 (3.2%)	0 (0%)			
Jumping from a height	4 (14.8%)	0 (0%)	2 (11.8%)	3 (15.8%)	0 (0%)	0 (0%)			
Firearm/sharp object	9 (33.3%)	0 (0%)	1 (5.9%)	3 (15.8%)	2 (6.5%)	4 (33.3%)			
Psychiatric consultation requested							$X^2=74.837$	$<0.001^{**}$	0.787
No	24 (88.9%)	4 (26.7%)	0 (0%)	1 (5.3%)	4 (12.9%)	9 (75%)			
Yes	3 (11.2%)	11 (73.3%)	17 (100%)	18 (94.7%)	27 (87.1%)	3 (25%)			

n: Number of patients, p: Significance level, X^2 : Chi-square statistic, *Fisher's exact test, **Pearson Chi-square test

Table 4. Comparison of ICU length of stay by method of suicide attempt and presence of physical illness (n=121)

	ICU stay (days) median (Q1-Q3)	Statistic	p	Effect size (η^2/r)
Method of suicide attempt		KW=10.138	0.102*	0.059
Drug ingestion (n=84; 69.4%)	2 (1-4)			
Hanging (n=9; 7.4%)	7 (1.5-8.5)			
Jumping from a height (n=9; 7.4%)	6 (1-15)			
Firearm/sharp object (n=19; 15.7%)	3 (1-6)			
Comorbid physical illness		Z=-0.855	0.393**	-0.078
No (n=94; 77.7%)	2 (1-5)			
Yes (n=27; 22.3%)	2 (1-5)			

KW: Kruskal-Wallis test statistic, p: Significance level, Q1: First quartile, Q3: Third quartile, Z: Mann-Whitney U test statistic, ICU: Intensive care unit, r: Rnk-biserial correlation, *Kruskal-Wallis test, **Mann-Whitney U test

Table 5. Univariate logistic regression analysis of factors predicting mortality following suicide attempts (0=survivor, 1=deceased) (n=121)

		B	SE	Wald	p	Unadjusted OR	95% CI
Age (years)		0.036	0.014	6.501	0.011	1.037	1.008-1.066
ICU stay duration (days)		0.100	0.044	5.038	0.025	1.105	1.013-1.206
Gender	Female =Ref.	3.343	1.040	10.339	0.001	28.311	3.689-217.269
Marital status	Married=Ref.	0.388	0.490	0.628	0.428	1.475	0.564-3.854
Residence	Urban=Ref.	0.651	0.457	2.023	0.155	1.917	0.782-4.698
Psychiatric diagnosis	No=Ref.	0.088	0.496	0.032	0.859	1.092	0.413-2.888
Psychotropic medication use	No=Ref.	-0.187	0.450	0.172	0.678	0.830	0.343-2.005
Comorbid physical illness	No=Ref.	1.209	0.478	6.395	0.011	3.352	1.313-8.557
History of suicide attempt	No=Ref.	-0.054	0.451	0.014	0.905	0.948	0.391-2.296
Psychiatric consultation	No=Ref.	-3.520	0.666	27.921	<0.001	0.030	0.008-0.109
Method of suicide attempt	Drug ingestion=Ref.						
Hanging		1.778	0.751	5.612	0.018	5.920	1.359-25.781
Jumping from a height		1.778	0.751	5.612	0.018	5.920	1.359-25.781
Firearm/sharp object		1.896	0.570	11.075	<0.001	6.660	2.180-20.345

B: Logit regression coefficient, SE: Standard error, p: Significance level, OR: Odds ratio, 95% CI: 95% confidence interval, R²: Nagelkerke R-squared value, ICU: Intensive care unit

was not significantly associated with the risk of mortality (all $p>0.05$). The variable of psychiatric consultation showed a strong inverse relationship with mortality. Not receiving a psychiatric consultation was significantly associated with increased risk of death ($p<0.001$). The model indicated that patients who received psychiatric consultation during their ICU stay were much less likely to die than those who did not (OR=0.03, 95% CI=0.008-0.109). The method of suicide attempt also emerged as a significant predictor. Using drug ingestion as the reference category, attempting suicide by hanging or jumping from a height increased the odds of death approximately 6-fold, while using firearms or sharp objects increased this risk by 6.7 times (corresponding p -values <0.05 -OR=5.920, 95% CI=1.359-25.781- and <0.001 -OR=6.660, 95% CI=2.180-20.345-, respectively). On the other hand, the year and season of the suicide attempt were not found to have a significant association with mortality in the multivariate model.

The results of the multivariable logistic regression analysis are presented in [Table 6](#). The overall model was statistically significant and demonstrated strong explanatory power

(Nagelkerke $R^2=0.772$). The area under the ROC curve indicated high discriminatory ability of the model (AUC: 0.964, 95% CI=0.921-1.000, $p<0.001$). The Hosmer-Lemeshow goodness-of-fit test was significant ($X^2=21.986$, $p=0.005$), suggesting that the model did not achieve perfect calibration. Age and ICU length of stay were not statistically significant predictors ($p>0.05$). Male gender was significantly associated with an increased risk of mortality compared to females (aOR=35.402, 95% CI=1.821-688.370). The presence of comorbid physical illness increased the risk of death, although this association did not reach statistical significance (aOR=10.953, 95% CI=0.651-184.183). Lack of psychiatric consultation was strongly and significantly associated with mortality (aOR=0.007, 95% CI=0.000-0.094). Regarding methods of suicide attempt, jumping from a height was independently associated with a markedly increased risk of death compared to drug ingestion (aOR=57.931, 95% CI=2.912-1152.486). Hanging (aOR=11.732, 95% CI=0.333-413.939) and firearm/sharp object use (aOR=1.912, 95% CI=0.314-11.630) showed elevated odds ratios but did not reach statistical significance.

Table 6. Multivariate logistic regression analysis of factors predicting mortality following suicide attempts (0=survivor, 1=deceased) (n=121)

		B	SE	Wald	p	Adjusted OR	95% CI
Age (years)		0.037	0.040	0.858	0.354	1.037	0.960-1.121
ICU stay duration (days)		0.087	0.074	1.382	0.240	1.090	0.944-1.260
Gender	Female=Ref.	3.567	1.514	5.549	0.018	35.402	1.821-688.370
Comorbid physical illness	No=Ref.	2.394	1.440	2.763	0.096	10.953	0.651-184.183
Psychiatric consultation	No=Ref.	-5.012	1.353	13.722	<0.001	0.007	0.000-0.094
Method of suicide attempt	Drug ingestion=Ref.						
Hanging		2.462	1.818	1.834	0.176	11.732	0.333-413.939
Jumping from a height		4.059	1.526	7.078	0.008	57.931	2.912-1152.486
Firearm/sharp object		0.648	0.921	0.495	0.482	1.912	0.314-11.630

Hosmer-Lemeshow: $X^2=21.986$, $p=0.005$, Nagelkerke $R^2=0.772$, AUC: 0.964 (95% CI=0.921-1.000, $p<0.001$)

B: Logit regression coefficient, SE: Standard error, p: Significance level, OR: Odds ratio, 95% CI: 95% confidence interval, R²: Nagelkerke R-squared value

DISCUSSION

The present study investigated the factors associated with mortality among patients monitored in intensive care units following suicide attempts. It also examined the clinical and demographic characteristics and the timing of attempts among individuals who engaged in highly lethal methods. According to the findings of the present study, advanced age, male gender, the presence of comorbid physical illnesses, use of highly lethal suicide methods, more extended ICU stay, and whether a psychiatric consultation was requested during the ICU process were all significantly associated with mortality. Patients who used highly lethal methods were most often either referred to closed psychiatric wards or died during hospitalization. Notably, male gender and the lack of psychiatric consultation during ICU stay were the variables most strongly associated with mortality.

In the present study, the mean age of patients who died following ICU admission was found to be significantly higher compared to those who survived. Consistent with the literature, older adults, particularly older males, have been shown to have a higher risk of death following suicide attempts.⁷ This has been attributed to the greater planning involved and the use of more lethal methods in this age group.³⁶ Additionally, the presence of more comorbid medical conditions and increased vulnerability to medical complications among older individuals may contribute to this elevated risk.¹⁶ Psychosocial factors such as increased social isolation, higher rates of spousal loss, and living alone are also thought to play a role in the increased suicide mortality observed in older adults.^{7,15}

The present study focused on severe suicide attempts and found that the majority of patients requiring ICU care due to suicide attempts were male. Consistent with our findings, previous studies have reported a higher incidence of serious suicide attempts among males.³⁷ Moreover, in our sample, male gender emerged as one of the variables most strongly associated with mortality following ICU admission. Similarly, the literature indicates that males are significantly more likely to die from suicide attempts compared to females.^{6,38} In a large-scale study, the fatality rate of suicide attempts was reported as 14.7% for males, compared to only 3.3% for females.³⁹ This discrepancy has been attributed to males' tendency to use more lethal methods and their generally stronger intent to die.^{40,41} Our findings support this pattern, demonstrating a higher rate of mortality among males during ICU follow-up and underscoring the prognostic significance of gender. These results suggest that males who attempt suicide constitute a particularly high-risk group requiring targeted preventive and clinical interventions.

In contrast to some prior studies, the present study did not find a significant association between mortality and marital status, seasons, or patients' place of residence. However, given that our study focused on high-lethality suicide attempts, it is noteworthy that the majority of individuals involved in these attempts were unmarried. Previous research has indicated that the lethality of suicide attempts is significantly higher among unmarried individuals.⁴² Although some studies have

reported that suicide attempts in rural areas tend to be more lethal, other research has found no significant differences based on place of residence.⁴³ Similarly, our analysis did not reveal a significant association between attempt lethality and residential setting. Regarding seasonal variation, earlier studies have shown that completed suicides tend to occur more frequently in the spring.^{20,23,24} While our findings did not demonstrate a statistically significant relationship between season and mortality, this may be attributable to the relatively low number of completed suicides in our sample. Nevertheless, given our study's focus on serious suicide attempts, it is notable that most of these attempts occurred during the spring months. Another key finding is that the majority of high-lethality suicide attempts in our sample occurred between 2021 and 2023. Although global data on yearly peaks in severe suicide attempts are limited, studies have shown a marked increase in serious suicide attempts and ICU admissions following the COVID-19 pandemic.⁴⁴ This temporal pattern aligns with the time frame observed in our study. Although we observed apparent post-pandemic variation in suicide attempts, we lacked data on total ICU admissions for each year. Therefore, denominators could not be provided. This limits our ability to determine whether changes reflect absolute increases or shifts relative to overall ICU utilization.

In terms of clinical characteristics, approximately 72.7% of the sample had a previously diagnosed psychiatric disorder. Prior studies have shown that nearly 90% of individuals who engage in serious suicide attempts requiring intensive medical intervention have at least one psychiatric disorder.¹⁰⁻¹² Furthermore, high-lethality attempts are often associated with the presence of multiple co-occurring psychiatric conditions.¹³ In the present study, no significant association was found between prior psychiatric diagnosis and mortality. This finding may be attributed to the relatively small number of completed suicides in our sample.

A prior suicide attempt is considered one of the most significant risk factors for serious suicidal behavior. In our sample, 38% of the patients had a history of at least one previous suicide attempt. Consistent with our findings, previous studies have reported that 20% to 40% of individuals who engage in serious suicide attempts have a history of prior attempts.^{45,46} In the present study, no significant difference was found between survivors and non-survivors in terms of prior suicide attempts. This may be attributed to the fact that all patients in the sample had engaged in high-lethality suicide attempts.

Previous research has demonstrated that individuals with physical illnesses are significantly more likely to engage in severe and lethal suicide attempts compared to those without such conditions.^{16,17} Consistent with the literature, our study also found that the rate of mortality following suicide attempts was significantly higher among individuals with comorbid physical illnesses. This finding may be explained by factors such as functional impairment associated with chronic physical diseases, the burden of managing long-term illness, co-occurring psychiatric conditions, and the increased lethality of suicide attempts when compounded by underlying physical health issues.

The method used in a suicide attempt is a crucial factor in determining the individual's survival outcome. In the present study, the majority of survivors (78.7%) attempted suicide by ingesting drugs or toxic substances. In contrast, methods such as hanging, jumping from heights, and the use of firearms or sharp instruments were more common among the fatal cases. According to the literature, while drug or poison ingestion accounts for approximately 60% of all suicide attempts, it is responsible for only about 13% of suicide deaths. Conversely, firearms and hanging constitute only around 9% of all attempts but account for more than 75% of completed suicides.³⁹ These findings are consistent with our results and highlight the importance of considering the method of suicide attempt as a critical risk indicator.

One of the key findings of the present study is that the mortality rate was significantly higher among patients who did not receive a psychiatric evaluation compared to those who did. Several factors may account for this observation. First, it is possible that the severity of the patient's medical condition precluded a psychiatric consultation. In cases where death occurred shortly after admission due to critical clinical status, there may not have been sufficient time or opportunity for psychiatric assessment. In this context, the lack of consultation could be seen as a consequence of poor prognosis. Alternatively, for patients who did undergo psychiatric evaluation, the interventions provided may have indirectly contributed positively to the intensive care process. For instance, psychological stabilization may enhance treatment adherence and reduce complications such as agitation during ICU stay.⁴⁷ Another possibility is that psychiatric consultations were more commonly requested for patients who were closer to discharge. Previous studies have reported that some patients admitted to the ICU due to suicide attempts do not receive psychiatric assessment, or that such evaluations are delayed.²⁹ This may lead to missed intervention opportunities, increased risk of repeated suicide attempts, and ultimately, higher mortality rates. Nevertheless, the observed consultation-mortality association likely reflects confounding by severity and immortal-time bias, rather than a direct protective and causal effect.

Another notable finding of the present study is the association between the type of suicide attempt, the request for psychiatric consultation, and the type of discharge from the intensive care unit. In our sample, patients who attempted suicide using more lethal methods, such as hanging, jumping from a height, or the use of firearms or sharp objects, were predominantly referred to psychiatric units, especially closed psychiatric wards. According to the literature, inpatient treatment in closed psychiatric settings is often recommended for individuals who have engaged in high-lethality suicide attempts, as it provides both enhanced safety and opportunities for more effective intervention.^{48,49} Patients who attempt serious suicide acts frequently present with significant psychiatric burden, agitation, aggression, and impulsivity, factors that often necessitate treatment in a closed psychiatric ward.⁴⁹ In contrast, some patients in our study were referred for outpatient psychiatric follow-up. It has been suggested that outpatient care may be appropriate for

individuals who do not exhibit active suicidal ideation, lack severe psychiatric symptoms, have adequate social support, and demonstrate good insight and treatment adherence following psychiatric evaluation.³⁵

In the present study, a substantial proportion of patients who were discharged against medical advice had attempted suicide using firearms or sharp/piercing objects. The literature indicates that this patient group typically consists of individuals with poor insight, low treatment adherence, high impulsivity, and a fear of stigmatization.³⁴

The study found that patients who attempted suicide by hanging or jumping from a height had more extended stays in the ICU. These high-lethality methods tend to prolong ICU treatment due to the complications and organ failures they cause. Previous studies have reported that ICU stays due to major trauma typically range from 9 to 17 days, whereas in cases of severe poisoning, the duration is generally between 6 and 10 days.^{50,51}

In the present study, the method of suicide attempt was strongly associated with mortality, with approximately a sixfold increased risk of death. Specifically, the use of firearms or sharp objects was associated with a 6.7-fold increased risk of death. These findings underscore that the method of suicide attempt may be a critical determinant of clinical course and prognosis. Similarly, previous research has demonstrated that suicide attempts by hanging increase the risk of death by 6.2 times and jumping from a height by 3.7 times.^{52,53} In suicide attempts involving sharp objects, targeting areas other than the wrist or forearm has been shown to elevate the risk of death by 1.7 to 4.5 times.^{54,55} Rapid identification of individuals who employ high-lethality methods, along with prompt medical intervention and close monitoring, is crucial for reducing mortality in this population.

Limitations

This study has several limitations. First, it employed a retrospective design, which limits the ability to establish causal relationships. The apparent association between the absence of psychiatric consultation and higher mortality may partly reflect early deaths or greater illness severity, rather than a direct effect. The data were collected from hospitals in a single province, which restricts the generalizability of the findings. The sample size was relatively small. The absence of long-term follow-up data limits prospective risk assessment. Additionally, detailed information on certain important risk factors, such as patients' socioeconomic status and family history, was not available. The study did not include a control group and only evaluated suicide attempts requiring intensive care. Since the data were retrospectively extracted from patient medical records, the risk of incomplete or inaccurate documentation, although carefully reviewed, cannot be entirely ruled out. Physical illness was defined based on documented chronic medical conditions, while acute conditions were not consistently recorded, which reduced the granularity of this variable. Firearm and sharp-object attempts were grouped together. Although these methods differ in lethality and clinical trajectory, the small number

of cases in each subgroup precluded separate analyses. This grouping may have obscured potential differences between these methods and represents another limitation of the study. Finally, our dataset did not include detailed timestamp information for psychiatric consultations or severity proxies such as mechanical ventilation, vasopressor use, or multi-organ failure indicators. Consequently, we were unable to refit models excluding early deaths, to model consultation as a time-dependent exposure, or to adjust for illness severity using APACHE II, SOFA, or comparable surrogate variables.

CONCLUSION

This retrospective study examined the clinical and demographic characteristics of patients admitted to the ICU following a suicide attempt and identified factors associated with suicide-related mortality. The findings suggest that older age, male gender, the presence of comorbid physical illnesses, the use of high-lethality suicide methods, and the absence of psychiatric consultation are significantly associated with increased mortality risk. Patients who employed high-lethality methods tended to have more extended ICU stays and were more frequently referred to closed psychiatric units. It has been revealed that timely psychiatric evaluation may be a critical factor in the clinical course of patients, and mortality rates may be significantly higher in patients who do not receive psychiatric consultation. This underscores the importance of psychiatric involvement in the ICU setting. Overall, the findings highlight the need for early identification of high-risk individuals, the development of appropriate intervention strategies, and improvements in post-attempt care processes. Future research should explore these risk factors in greater detail and focus on developing effective prevention and management strategies for individuals at risk of suicide.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Giresun Training and Research Hospital Ethics Committee (Date: 23.07.2025, Decision No: 23.07.2025/14).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Goodfellow B, Kolves K, De Leo D. Contemporary definitions of suicidal behavior: a systematic literature review. *Suicid Life-Threat Behav.* 2019;49(2):488-504. doi:10.1111/sltb.12457
- Pirkis J, Dandona R, Silverman M, Khan M, Hawton K. Preventing suicide: a public health approach to a global problem. *Lancet Public Health.* 2024;9(10):e787-e795. doi:10.1016/S2468-2667(24)00149-X
- Weekly CC. World suicide prevention day-September 10, 2020. doi:10.46234/ccdcw2020.191
- World Health Organization. Suicide worldwide in 2021: global health estimates. Geneva, Switzerland: World Health Organization; 2025.
- Klonsky ED, May AM, Saffer BY. Suicide, suicide attempts, and suicidal ideation. *Ann Rev Clin Psychol.* 2016;12(1):307-330. doi:10.1146/annurev-clinpsy-021815-093204
- Glenn CR, Kleiman EM, Kellerman J, et al. Annual research review: a meta-analytic review of worldwide suicide rates in adolescents. *J Child Psychol Psychiatr.* 2020;61(3):294-308. doi:10.1111/jcpp.13106
- Beghi M, Butera E, Cerri CG, et al. Suicidal behaviour in older age: A systematic review of risk factors associated to suicide attempts and completed suicides. *Neurosci Biobehav Rev.* 2021;127:193-211. doi:10.1016/j.neubiorev.2021.04.011
- Bachmann S. Epidemiology of suicide and the psychiatric perspective. *Int J Environment Res Public Health.* 2018;15(7):1425. doi:10.3390/ijerph15071425
- Kim SH, Kim HJ, Oh SH, Cha K. Analysis of attempted suicide episodes presenting to the emergency department: comparison of young, middle aged and older people. *Int J Mental Health Systems.* 2020;14:1-8. doi:10.1186/s13033-020-00378-3
- Hoertel N, Franco S, Wall MM, et al. Mental disorders and risk of suicide attempt: a national prospective study. *Mol Psychiatr.* 2015;20(6):718-726. doi: 10.1038/mp.2015.19
- Rizk MM, Galfalvy H, Miller JM, et al. Characteristics of depressed suicide attempters with remitted substance use disorders. *J Psychiatr Res.* 2021;137:572-578. doi:10.1016/j.jpsychires.2020.10.041
- Ivey-Stephenson AZ, Crosby AE, Hoenig JM, et al. Suicidal thoughts and behaviors among adults aged ≥18 years-United States, 2015-2019. *MMWR Surveill Summ.* 2022;71(1):1-19. doi:10.15585/mmwr.ss7101a1
- Xu YE, Barron DA, Sudol K, Zisook S, Oquendo MA. Suicidal behavior across a broad range of psychiatric disorders. *Mol Psychiatr.* 2023;28(7):2764-2810. doi:10.1038/s41380-022-01935-7
- Amiri S. Unemployment and suicide mortality, suicide attempts, and suicide ideation: a meta-analysis. *Int J Mental Health.* 2022;51(4):294-318. doi: 10.1080/00207411.2020.1859347
- Øien-Ødegaard C, Hauge LJ, Reneflot A. Marital status, educational attainment, and suicide risk: a Norwegian register-based population study. *Populat Health Metrics.* 2021;19:1-11. doi: 10.1186/s12963-021-00263-2
- Torino G, Rignanese M, Salmè E, et al. Physical pain and suicide-related outcomes across the lifespan: systematic review and meta-analysis. *Psychiatr Res.* 2025;116371. doi:10.1016/j.psychres.2025.116371
- Onyeka IN, Maguire A, Ross E, O'Reilly D. Does physical ill-health increase the risk of suicide? A census-based follow-up study of over 1 million people. *Epidemiol Psychiatr Sci.* 2020;29:e140. doi:10.1017/s2045796020000529
- Prazak M, Bacigalupi R, Hamilton SC. Rural suicide: demographics, causes, and treatment implications. *Comm Ment Health J.* 2025;61(1):66-75. doi:10.1007/s10597-024-01327-x
- Li L, Liu C, Tong Y, et al. Comparison of characteristics of suicide attempts registered in urban and rural areas in China. *Front Psychiatr.* 2022;12:805324. doi:10.3389/fpsyt.2021.805324
- Ambar Akkaoui M, Chan-Chee C, Laaidi K, et al. Seasonal changes and decrease of suicides and suicide attempts in France over the last 10 years. *Scientific Reports.* 2022;12(1):8231. doi:10.1038/s41598-022-12215-3
- Su MK, Chan PY, Hoffman RS. The seasonality of suicide attempts: a single poison control center perspective. *Clin Toxicol (Phila).* 2020; 58(11):1034-1041. doi:10.1080/15563650.2020.1733591
- Tifha M, Abbes W, Mahdhaoui W, et al. Effect of seasonality, climatic and temporal factors on suicide attempts amongst patients from southern tunisia. *Eur Psychiatr.* 2021;64(S1):S587-S587. doi:10.1192/j.eurpsy.2021.1566

23. Coimbra DG, Pereira ESAC, de Sousa-Rodrigues CF, et al. Do suicide attempts occur more frequently in the spring too? A systematic review and rhythmic analysis. *J Affect Disord.* 2016;196:125-37. doi:10.1016/j.jad.2016.02.036
24. Akkaya-Kalayci T, Vyssoki B, Winkler D, et al. The effect of seasonal changes and climatic factors on suicide attempts of young people. *BMC Psychiatry.* 2017;17(1):365. doi:10.1186/s12888-017-1532-7
25. Quesada-Franco M, Braquehais MD, Valero S, et al. A comparison of medically serious suicide attempters admitted to intensive care units versus other medically serious suicide attempters. *BMC Psychiatry.* 2022;22(1):805. doi:10.1186/s12888-022-04427-8
26. Renuka MK, Kalaiselvan MS, Arunkumar AS. An analysis of the predictors of mortality and morbidity in patients admitted after suicidal hanging to an Indian multidisciplinary intensive care unit. *Indian J Anaesth.* 2017;61(7):538-542. doi:10.4103/ija.IJA_170_17
27. Hirot F, Aminata A, Azouvi P, et al. Factors influencing the length of stay in a unit with mixed rehabilitation and psychiatric care after a suicide attempt by jumping: a cohort observational retrospective study. *Eur J Physic Rehabil Med.* 2022;58(5):749. doi:10.23736/S1973-9087.22.07341-5
28. El-Qawaqzeh K, Colosimo C, Bhogadi SK, et al. Unequal treatment? Confronting racial, ethnic, and socioeconomic disparity in management of survivors of violent suicide attempt. *J Am College Surg.* 2023;237(1):68-78. doi:10.1097/XCS.0000000000000716
29. Marwale AV, Phadke SS, Kocher AS. Psychiatric management of patients in intensive care units. *Indian J Psychiatry.* 2022;64(Suppl 2):S292-S307. doi:10.4103/indianjpsychiatry.indianjpsychiatry_28_22
30. Chung CH, Chien WC, Yeh HW, Tzeng NS. Psychiatric consultations as a modifiable factor for repeated suicide attempt-related hospitalizations: a nationwide, population-based study. *J Affect Disord.* 2021;278:157-164. doi:10.1016/j.jad.2020.09.013
31. Son JM, Jeong J, Ro YS, et al. Effect of previous emergency psychiatric consultation on suicide re-attempts-a multi-center observational study. *Am J Emerg Med.* 2020;38(9):1743-1747. doi:10.1016/j.ajem.2020.05.030
32. Kim H, Kim Y, Shin M-H, et al. P68: Early psychiatric referral after attempted suicide helps prevent suicide reattempts: a longitudinal national cohort study in South Korea. *Int Psychogeriatr.* 2023;35(S1):244-245. doi:10.1017/S1041610223004064
33. Meyer HL, Reck T, Polan C, et al. A Comparison of severely injured patients after suicide attempts and violent crimes-a retrospective study of a level 1 trauma center. *Clin Pract.* 2024;14(4):1468-1477. doi:10.3390/clinpract14040118
34. Gilgoff JN, Wagner F, Frey JJ, Osteen PJ. Help-seeking and man therapy: the impact of an online suicide intervention. *Suicid Life Threat Behav.* 2023;53(1):154-162. doi:10.1111/sltb.12929
35. Che SE, Gwon YG, Kim KH. Follow-up timing after discharge and suicide risk among patients hospitalized with psychiatric illness. *JAMA Network Open.* 2023;6(10):e2336767-e2336767. doi:10.1001/jamanetworkopen.2023.36767
36. Kim DW, Cho SE, Kang JM, et al. Risk factors for serious suicide attempts: difference between older and younger attempters in the emergency department. *Front Psychiatr.* 2021;11:607811. doi:10.3389/fpsy.2020.607811
37. Freeman A, Mergl R, Kohls E, et al. A cross-national study on gender differences in suicide intent. *BMC Psychiatry.* 2017;17(1):234. doi:10.1186/s12888-017-1398-8
38. Tai HC, Huang SH, Huang YC, et al. Gender differences in death during hospitalization following a survived suicide attempt in Taiwan from 1998 to 2015: a cross-sectional study. *J Men's Health.* 2022;18(4):103. doi:10.31083/j.jomh1804103
39. Conner A, Azrael D, Miller M. Suicide case-fatality rates in the United States, 2007 to 2014: a nationwide population-based study. *Ann Int Med.* 2019;171(12):885-895. doi:10.7326/M19-1324
40. Barry R, Rehm J, de Oliveira C, Gozdyra P, Chen S, Kurdyak P. Rurality as a risk factor for attempted suicide and death by suicide in Ontario, Canada. *Canad J Psychiatr.* 2022;67(9):679-689. doi:10.1177/07067437211053300
41. Mehanović E, Rosso G, Cuomo GL, et al. Risk factors for mortality after hospitalization for suicide attempt: results of 11-year follow-up study in Piedmont Region, Italy. *Soc Psychiatr Psychiatr Epidemiol.* 2024;59(6):1039-1051. doi:10.1007/s00127-023-02544-7
42. Akbarizadeh F, Hajivandi A, Hajivandi M, Zeidabadinejad MS. Marriage related suicide fatality rates. *Iran J Psychiatry.* 2019;14(1):54-59.
43. Barry R, Rehm J, de Oliveira C, Gozdyra P, Kurdyak P. Rurality and risk of suicide attempts and death by suicide among people living in four English-speaking high-income countries: a systematic review and meta-analysis. *Canad J Psychiatr.* 2020;65(7):441-447. doi:10.1177/0706743720902655
44. Beneria A, Marte L, Quesada-Franco M, et al. Trends in medically serious suicide attempts before and after COVID-19: a four-year retrospective analysis (2018-2022). *BMC Psychiatry.* 2024;24(1):770. doi:10.1186/s12888-024-06232-x
45. Kim MJ, Yang JH, Koh MJ, Kim Y, Lee B, Ahn YM. Risk factors of reattempt among suicide attempters in South Korea: a nationwide retrospective cohort study. *PLoS One.* 2024;19(4):e0300054. doi:10.1371/journal.pone.0300054
46. Goñi-Sarriés A, Blanco M, Azcárate L, Peinado R, López-Goñi JJ. Are previous suicide attempts a risk factor for completed suicide? *Psicothema.* 2018;30(1):33-38. doi:10.7334/psicothema2016.318
47. Blair KTA, Eccleston SD, Binder HM, McCarthy MS. Improving the patient experience by implementing an ICU diary for those at risk of post-intensive care syndrome. *Journal of patient experience.* 2017;4(1):4-9. doi:https://doi.org/10.1177/2374373517692927
48. Berardelli I, Innamorati M, Sarubbi S, et al. Demographic and clinical correlates of high-lethality suicide attempts: a retrospective study in psychiatric inpatients. *J Psychiatr Pract.* 2021;27(6):410-416. doi:10.1097/PRA.0000000000000579
49. Gvion Y, Levi-Belz Y. Serious suicide attempts: systematic review of psychological risk factors. *Front Psychiatr.* 2018;9:56. doi:10.3389/fpsy.2018.00056
50. Ono Y, Ishida T, Tomita N, et al. Attempted suicide is independently associated with increased in-hospital mortality and hospital length of stay among injured patients at community tertiary hospital in Japan: a retrospective study with propensity score matching analysis. *Int J Environ Res Public Health.* 2024;21(2):121. doi:10.3390/ijerph21020121
51. Nohl A, Ohmann T, Kamp O, et al. Major trauma due to suicide attempt: increased workload but not mortality. *Eur J Trauma Emerg Surg.* 2022;48(1):519-523. doi:10.1007/s00068-020-01436-8
52. Runeson B, Tidemalm D, Dahlin M, Lichtenstein P, Långström N. Method of attempted suicide as predictor of subsequent successful suicide: national long term cohort study. *BMJ.* 2010;341:c3222. doi:10.1136/bmj.c3222
53. Lee SY, Song KJ, Hong KJ, et al. Epidemiology and outcomes of severe injury patients related to suicide attempt or self-harm in Korea: nationwide community-based study, 2016-2020. *Injury.* 2025;56(2):112032. doi:10.1016/j.injury.2024.112032
54. Geulayov G, Casey D, Bale E, et al. Risk of suicide in patients who present to hospital after self-cutting according to site of injury: findings from the multicentre study of self-harm in England. *Psychol Med.* 2023;53(4):1400-1408. doi:10.1017/S0033291721002956
55. Carroll R, Thomas KH, Bramley K, et al. Self-cutting and risk of subsequent suicide. *J Affect Disord.* 2016;192:8-10. doi:10.1016/j.jad.2015.12.007

An underrecognized aspect of oncology care: temporomandibular joint evaluation in patients with head and neck cancers

 Melisa Öçbe¹,  Mahmut Sabri Medişoğlu²,  Çetin Batuhan Öçbe³

¹Department of Oral and Maxillofacial Radiology, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Türkiye

²Department of Anatomy, Faculty of Dentistry, Kocaeli Health and Technology University, Kocaeli, Türkiye

³Department of Family Medicine, Ataşehir Family Medicine Center, Kartepe District Health Directorate, Ministry of Health, Kocaeli, Türkiye

Cite this article as: Öçbe M, Medişoğlu MS, Öçbe ÇB. An underrecognized aspect of oncology care: temporomandibular joint evaluation in patients with head and neck cancers. *J Health Sci Med.* 2025;8(6):1143-1148.

Received: 16.09.2025

Accepted: 14.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to evaluate and compare temporomandibular joint (TMJ) disc position and condylar morphology on magnetic resonance imaging (MRI) and computed tomography (CT) images of patients with head and neck cancer (HNC) and healthy controls, to determine the prevalence and coexistence of soft-tissue and osseous alterations and to explore their potential clinical relevance.

Methods: In this retrospective case-control study, 145 HNC patients and 147 controls that underwent standardized maxillofacial CT and MRI. Anterior disc displacement (ADD) was defined on mid-sagittal MRI; condylar morphology (normal, flattening, erosion, osteophyte) was classified on sagittal CT.

Results: Two hundred ninety-two individuals were analyzed (HNC 49.7%). Age was similar (mean difference -3.11 years; $p=0.097$; Mann-Whitney $p=0.0737$). ADD prevalence was higher in HNC: right 49.0% vs 37.4% (RD +11.6 percentage points [pp]; RR 1.31; $\chi^2 p=0.046$), left 52.4% vs 32.7% (RD +19.8 pp; RR 1.61; $\chi^2 p=0.00064$), and any ADD 60.7% vs 46.9% (RD +13.8 pp; RR 1.29; $\chi^2 p=0.018$); significance was retained after Holm-Bonferroni adjustment. Condylar morphology distributions differed on both sides (right $\chi^2 p=9.6 \times 10^{-5}$; left $p=8.1 \times 10^{-10}$), driven by more erosion in HNC (right 26.9% vs 11.6%; left 27.6% vs 13.6%). “Any condylar pathology” was more frequent in HNC (93.8% vs 85.7%; RD +8.1 pp; OR 2.52; $\chi^2 p=0.023$; Fisher $p=0.033$). ADD showed no lateral preference within groups (McNemar $p>0.30$). Controls, but not HNC patients, exhibited right-greater-than-left discordance for “any condylar pathology” (McNemar $p=2.5 \times 10^{-5}$ vs 1.000).

Conclusion: HNC is associated with a high burden of TMJ abnormalities; particularly ADD and an erosive change in condyle; consistent with treatment- and disease-related vulnerability of joint tissues. Providing TMJ assessment into HNC imaging, coupled with early jaw-mobilization and masticatory-muscle-sparing strategies, may help preserve function.

Keywords: Temporomandibular joint, head and neck neoplasms, magnetic resonance imaging, tomography, X-Ray computed, condyle, neoplasm-associated inflammation

INTRODUCTION

Head and neck cancer (HNC) is managed with various oncologic therapies; including surgery, radiotherapy (external beam and brachytherapy), and systemic agents; which, while improving survival, are associated with substantial acute and late adverse effects.¹⁻⁷ Beyond well-recognized oral toxicities such as mucositis, xerostomia, dysgeusia, dysphagia, and increased caries risk, both the disease and its treatments can affect the stomatognathic system.⁶ Fibrosis of the masticatory muscles, radiotherapy-related trismus, postsurgical scarring, prolonged immobilization, and treatment-related immunosuppression may collectively impair jaw function and tissue homeostasis.⁷⁻⁹

Temporomandibular disorders (TMD) encompass pain, limited mouth opening, and functional limitations that

negatively impact quality of life-an especially critical consideration in patients already burdened by cancer care. Structurally, TMD includes degenerative osseous remodeling of the mandibular condyle (e.g., flattening, erosion, osteophyte formation) and internal derangements such as anterior disc displacement (ADD) with or without reduction.^{10,11} Despite their clinical importance, temporomandibular joint (TMJ) alterations in HNC populations remain under-characterized, and it is unclear to what extent they reflect pre-existing susceptibility versus treatment- or disease-related changes.

Imaging plays a central role in this context. Computed tomography (CT) optimally depicts osseous morphology and cortical integrity, whereas magnetic resonance imaging (MRI) is the reference standard for evaluating disc position

Corresponding Author: Melisa Öçbe, melisa.ocbe@kocaelisaglik.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

and soft-tissue structures.^{12,13} Accordingly, combined CT and MRI provide a comprehensive framework to quantify TMJ structural changes in patients undergoing HNC management.

This study aimed to compare the prevalence and pattern of TMJ abnormalities (MRI-defined disc displacement and CT-defined condylar morphology) between patients with HNC and age- and sex-matched controls. We hypothesized that HNC patients would exhibit a higher burden of disc displacement and degenerative condylar changes, with exploratory analyses assessing side predominance and within-person asymmetry.

METHODS

Ethical approval of this study was granted from the Kocaeli Health and Technology University Non-interventional Researches Ethics Committee (Date: 10.03.2025, Decision No: 2025-128). Only patients with informed consent form were included. All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

The study group included 145 patients with histopathologically confirmed head and neck malignancies: nasopharyngeal carcinoma (n=87), oral squamous cell carcinoma (n=21), sinonasal carcinoma (squamous cell carcinoma or adenocarcinoma; n=13), salivary gland malignancies (n=4), HPV-associated oropharyngeal carcinoma (n=1), basal cell carcinoma of ear (n=1), skull-base carcinomas (n=7), and metastatic lesions (n=11). The control group comprised 147 age- and sex-matched individuals who underwent imaging for non-neoplastic diagnostic purposes, with signed informed consent. Inclusion criteria for HNC were; histopathologically confirmed head and neck malignancy, initial cancer diagnosis made ≥ 6 months before the index imaging (to allow assessment of potential osseous changes), availability of both maxillofacial CT and TMJ MRI obtained using the standardized protocols described, performed within 90 days of each other. The maximum interval of 90 days between CT and MRI was selected to ensure both modalities represented the same disease and treatment phase while allowing adequate time for osseous remodeling to become radiographically apparent. Bone is a dynamic organ that undergoes continuous modeling and remodeling throughout life; in adults, resorption occurs over approximately 2-3 weeks, followed by a formation and mineralization phase lasting up to 3-4 months. These sequential processes underlie the time-dependent visibility of structural alterations in bone: soft-tissue and marrow changes appear early on MRI, whereas cortical erosions or surface remodeling become evident on CT only after several weeks of sustained osteoclastic activity. Therefore, the ≤ 90 -day window balances biologic plausibility with clinical feasibility, minimizing temporal bias while capturing both early soft-tissue and delayed osseous manifestations of TMJ changes in oncologic patients.¹⁴

Inclusion criteria for control group were; age- and sex-matched individuals with no current or prior head and neck malignancy, underwent maxillofacial CT and TMJ MRI for non-neoplastic diagnostic purposes, acquired with the same standardized protocols, within 90 days of each other, signed informed consent. Exclusion criteria for both HNC and

control group were; inadequate imaging quality or coverage precluding assessment (severe motion, metal artifacts, incomplete TMJ coverage, slice thickness or sequence parameters outside protocol), missing either modality (no CT or no TMJ MRI) or imaging performed >90 days apart, prior surgery directly altering TMJ anatomy (e.g., condylectomy, arthroplasty, mandibular reconstruction involving the condyle, segmental mandibulectomy affecting the condylar unit), maxillofacial trauma involving the mandible/TMJ, known inflammatory arthritides or systemic connective-tissue diseases primarily affecting the TMJ (e.g., rheumatoid arthritis, psoriatic arthritis) unless documented inactive and without TMJ involvement on imaging, congenital craniofacial syndromes or severe developmental anomalies impacting TMJ morphology (e.g., hemifacial microsomia) that preclude standard classification, prior radiotherapy confined to the TMJ region that obliterates anatomical landmarks, active osteonecrosis of the jaw with condylar destruction preventing classification, incomplete demographic data (age or sex) or withdrawal of consent.

Images from Kocaeli City Hospital, Department of Radiology archive were evaluated. CT images were obtained with a GE Revolution™ EVO/Optima™ CT660 scanner using standardized maxillofacial protocols (120 kVp, auto mA 100-250 mA, 0.625 mm slice thickness, 0.5 mm reconstruction interval, 250 mm FOV, 512×512 matrix). Multiplanar reformation in axial, coronal, and sagittal planes was performed for all CT datasets. MRI examinations were performed on a 3.0 T GE SIGNA™ Pioneer system with a head-and-neck coil and standardized parameters to ensure reproducibility. Axial T2 FSE, coronal T2 FSE, and coronal T2 STIR sequences were acquired with a 240×240 mm FOV, 512×512 matrix, 3 mm slice thickness, and 0.5 mm interslice gap. The imaging methods included CT and MRI of patients diagnosed with head and neck carcinomas, going through the treatment process.

All examinations were reviewed on a PACS workstation using bone windows for CT (typical W/L $\approx 1600/300$ HU) and routine soft-tissue windows for MRI. Sagittal oblique images were reconstructed perpendicular to the long axis of the mandibular condyle; coronal planes were parallel to it. Assessments were performed in maximum intercuspation (closed-mouth). The 12-o'clock position was defined at the superior apex of the condylar head. The posterior band of the disc was localized relative to this reference on the mid-condylar sagittal slice. Posterior band centered between approximately 11:30 and 12:30 (± 10 -15° of 12-o'clock), with the intermediate zone interposed between the condyle and articular eminence. Posterior band located anterior to $\sim 11:30$ (≥ 10 -15° anterior to 12-o'clock), or the intermediate zone not interposed between the condyle and eminence in closed-mouth images (**Figure 1**).

Morphology was classified on mid-sagittal sections (**Figure 2**). When multiple features coexisted, the most pathologic single category was assigned using the following hierarchy: osteophyte > erosion > flattening > normal. Normal: Smooth, ovoid condylar contour with preserved cortical outline and expected convex superior surface. Flattening: Loss of normal convexity with a planar/shelf-like superior surface or generalized rounding, with intact cortex; subcortical

sclerosis may be present but does not, by itself, advance the category. Erosion: Focal irregularity or interruption of the cortical margin with loss of condylar volume and/or subcortical cavitation compatible with active surface breakdown. Osteophyte: Focal exophytic marginal bony spur projecting beyond the expected condylar contour (commonly anterosuperior/superolateral), with otherwise preserved cortex at the spur base.

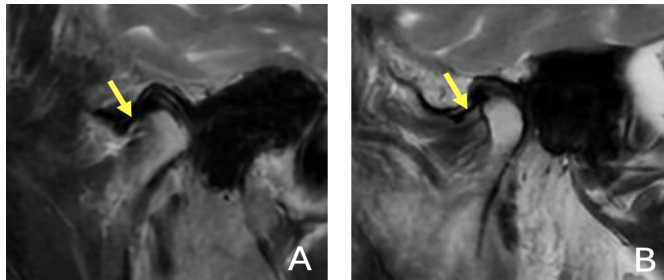


Figure 1. Right T2 weighted sagittal MR images showing the absence (A) and presence of anterior disc displacement (B) Anterior disc displacement

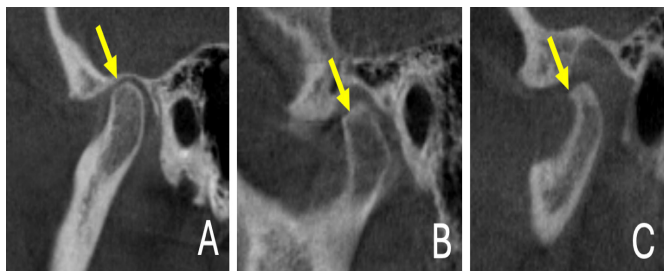


Figure 2. Sagittal computed tomography sections of the temporomandibular joint (bone algorithm). (A) Normal condyle: smooth, ovoid articular surface with preserved cortical outline and convex superior contour. (B) Condylar flattening: loss of normal convexity with a planar/shelf-like superior surface; subcortical sclerosis may be observed. (C) Osteophyte: focal exophytic marginal bony spur projecting beyond the expected condylar contour, typically at the antero- or superolateral margin.

Statistical Analysis

Analyses were performed in Python 3.11 using Pandas, SciPy, and Statsmodels. Analyses were two-sided ($\alpha=0.05$). Age was summarized as mean $\pm$ SD (median, IQR); normality (Shapiro-Wilk, Q-Q plots) and variance equality (Levene) guided use of Welch's t-test or Mann-Whitney U for group comparisons. Gender and binary outcomes were compared with Pearson's χ^2 (or Fisher's exact when sparse). Ordinal morphology distributions were compared with χ^2 tests of independence; trend (Cochran-Armitage) and proportional-odds logistic regression were planned sensitivities.

Within-person right-left differences used McNemar's test for paired binaries; Stuart-Maxwell/Bowker for paired ordinal data (or dichotomized McNemar if sparse). Effect sizes were reported: mean differences, risk differences, risk ratios, and odds ratios with 95% CIs (Newcombe/Wilson or exact).

Multiple testing within outcome families was controlled by Holm-Bonferroni. Residual age/sex imbalance was assessed (standardized mean differences) and addressed via multivariable logistic regression.

RESULTS

A total of 292 individuals were analyzed: 145 with head and neck cancer (HNC; 49.7%) and 147 age- and sex-matched controls (50.3%). Mean age did not differ significantly between groups (HNC 53.23 years vs. control 56.35 years; mean difference -3.11 years, 95% CI -6.79 to 0.57; Welch's $t=0.097$; Mann-Whitney U $p=0.0737$). Overall, 124/292 (42.5%) were female and 168/292 (57.5%) were male. Sex distribution was comparable by design (matched).

Although a few isolated cases of lateral or posterior disc displacement were observed, they were excluded from the analysis to maintain group homogeneity and ensure consistent diagnostic assessment. Consequently, only ADD was evaluated, as it constitutes the most prevalent and diagnostically reliable pattern observed on sagittal oblique MRI sections. Overall, right and left ADD were present in 126/292 (43.2%) and 124/292 (42.5%), respectively; 157/292 (53.8%) had ADD on either side (right-only 11.3%, left-only 10.6%, bilateral 31.8%).

Compared with controls, the HNC group showed higher ADD prevalence (Table 1). The between-group difference was largest for the left side (risk difference +19.8 percentage points [pp], RR 1.61, OR 2.27; $\chi^2 p=0.00064$). "Any ADD (right or left)" was also higher in HNC (risk difference +13.8 pp; RR 1.29; OR 1.75; $\chi^2 p=0.018$). Holm-Bonferroni correction across the three ADD comparisons (right, left, any) retained statistical significance for all χ^2 tests.

Any ADD was similar by sex (female 53.2% vs male 54.2%). Any condylar pathology was 94.4% in females and 86.3% in males. Mean age was 59.13 years in females and 51.61 years in males (descriptive, not powered for sex-effect inference). Within-person side comparisons showed no evidence of lateral preference for ADD in either group (HNC discordant R+/L-=12 vs R-/L+=17, McNemar $p=0.458$; controls 21 vs 14, $p=0.311$).

Across the cohort, right-side morphology was: normal 13.7%, flattening 63.4%, erosion 19.2%, osteophyte 3.8%; left-side: normal 23.3%, flattening 51.7%, erosion 20.5%, osteophyte 4.5%. Groupwise distributions differed markedly on both sides (Table 2). HNC patients exhibited substantially more erosions than controls (right: 26.9% vs 11.6%; left: 27.6% vs 13.6%), with overall $\chi^2 p=9.6\times 10^{-5}$ (right) and 8.1×10^{-10} (left). "Any condylar pathology" (either side $\neq$ normal) was more frequent in HNC than controls (93.8% vs 85.7%; risk difference +8.1

Table 1. Disc displacement prevalence by group with effect sizes

Outcome	Control n/N (%)	HNC n/N (%)	Risk difference, pp (HNC-Ctrl)	RR (95% CI)	OR (95% CI)	$\chi^2 p$	Fisher p
Right ADD	55/147 (37.4)	71/145 (49.0)	+11.6	1.31 (1.00-1.71)	1.60 (1.01-2.56)	0.046	0.059
Left ADD	48/147 (32.7)	76/145 (52.4)	+19.8	1.61 (1.21-2.12)	2.27 (1.41-3.65)	0.00064	0.00088
Any ADD (R or L)	69/147 (46.9)	88/145 (60.7)	+13.8	1.29 (1.04-1.60)	1.75 (1.10-2.78)	0.018	-

HNC: Head and neck cancer, pp: percentage points, RR: Risk ratio, OR: Odds ratio, ADD: Anterior disc displacement. Effect sizes reported with 95% CIs (Newcombe/Wilson or exact as applicable). Two-sided $\alpha=0.05$. Familywise multiplicity for ADD comparisons controlled by Holm-Bonferroni.

Table 2. Condylar morphology distribution by group and side

Side&category	Control n/N (%)	HNC n/N (%)	χ^2 p (4-level distribution)
Right			9.6×10 ⁻⁵
Normal	29/147 (19.7)	11/145 (7.6)	
Flattening	92/147 (62.6)	93/145 (64.1)	
Erosion	17/147 (11.6)	39/145 (26.9)	
Osteophyte	9/147 (6.1)	2/145 (1.4)	
Left			8.1×10 ⁻¹⁰
Normal	57/147 (38.8)	11/145 (7.6)	
Flattening	61/147 (41.5)	90/145 (62.1)	
Erosion	20/147 (13.6)	40/145 (27.6)	
Osteophyte	9/147 (6.1)	4/145 (2.8)	
Any condylar pathology (either side≠normal)	126/147 (85.7)	136/145 (93.8)	χ^2 p=0.023; Fisher p=0.033
χ^2 =Pearson's Chi-square. Holm-Bonferroni across the three χ^2 tests left all comparisons significant at $\alpha=0.05$			

pp, RR 1.09, OR 2.52; χ^2 p=0.023; Fisher p=0.033). Within-person asymmetry in “any condylar pathology” was evident in controls (discordant R+/L-=36 vs R-/L+=8; McNemar p=2.5×10⁻⁵), indicating more right-than-left involvement, whereas HNC patients were balanced (2 vs 2; p=1.000).

Coexisting disc displacement and condylar alterations were significantly more frequent among head and neck cancer patients (60.7%) compared with controls (45.6%) (χ^2 =6.12, p=0.013). This suggests that TMJ involvement in the oncologic population is more complex, often combining both soft-tissue and osseous changes (Table 3).

Table 3. Coexistence of disc displacement and condylar pathology by group

Group	n	Disc displacement (%)	Condylar pathology (%)	Both findings (%)
Control	147	46.9% (n=69)	85.7% (n=126)	45.6% (n=67)
HNC	145	60.7% (n=88)	93.8% (n=136)	60.7% (n=88)

HNC: Head and neck cancer

DISCUSSION

In this retrospective case-control study, patients with HNC exhibited a higher burden of (TMJ) alterations than controls across both soft-tissue and osseous domains. On MRI, HNC was associated with a significantly greater prevalence of ADD, most commonly found on the left side (risk difference +19.8 percentage points; RR 1.61; OR 2.27), and evident for the “any ADD” composite (+13.8 pp; RR 1.29; OR 1.75). HNC patients displayed significantly different condylar morphology distributions on both sides driven by increased erosions (right 26.9% vs 11.6%; left 27.6% vs 13.6%), and a higher overall prevalence of “any condylar pathology” (+8.1 pp; OR 2.52). In addition, coexistence of disc displacement and condylar changes is significantly more frequent in the head and neck cancer group than in the control group. These patterns remained significant after multiplicity control.

The increased prevalence of ADD and condylar pathology in HNC group can be multifactorial. For instance, radiotherapy

in head and neck area can induce hypovascular, hypocellular, and hypoxic bone and soft-tissue states; chemotherapy and systemic therapies may compound microcirculatory compromise and myofascial dysfunction. Muscle fibrosis and reduced range of motion can alter joint loading, promoting internal derangement and surface breakdown rather than slowly adaptive osteophyte formation.^{15,16} According to the study of Somay et al.,¹⁶ trismus in HNC appears tightly linked to masseter muscle volume, which functions as a pragmatic surrogate for treatment-related sarcopenia and localized muscle atrophy of the masticatory apparatus. Clinically, reduced masseter cross-sectional area not only signals higher trismus risk but also captures the broader catabolic state (cancer-associated sarcopenia) that undermines rehabilitation. For patients, preserving mouth opening is critical to enable adequate nutrition, speech, oral hygiene, and routine dental care,¹⁵⁻¹⁷ therefore, prevention should combine plan optimization.

In cancer patients, parafunction/bruxism due to oncologic stress, and nutritional changes (altered chewing patterns) can create eccentric loading and discal malposition.¹⁸ Additionally, tumor location and nodal disease might secondarily affect masticatory mechanics even without direct TMJ invasion.^{19,20} The relative excess of erosion and the relative paucity of osteophytes in HNC (compared with controls) is consistent with an active catabolic process with insufficient osteoblastic compensation.

An interesting divergence emerged in within-person asymmetry. Controls showed right-greater-than-left discordance for “any condylar pathology,” whereas HNC patients were balanced. ADD displayed no side preference in either cohort. The control pattern may reflect habitual chewing dominance or referral/scan-indication bias, while the side-symmetry in HNC hints at systemic or generalized drivers (treatment effects, systemic inflammation) rather than focal, side-specific insults. To apply these findings in clinical settings, as CT and MRI are already acquired for oncologic reasons, a structured TMJ read (disc position on MRI; condylar surface on CT) may add minimal burden yet may identify patients at risk for trismus, pain, or functional decline. Referral for jaw-mobilization programs, therapist-guided stretching, and individualized exercise splints may prevent progression from displacement to chronic dysfunction. Close collaboration among oral and maxillofacial radiology, radiation oncology, head-and-neck surgery, dentistry, and rehabilitation medicine can convert incidental TMJ findings into actionable, patient-centered plans.

Importantly, the results of this study reveal a clear pattern of increased TMJ involvement among patients with HNC. This may represents a meaningful clinical implications. The higher frequency of both disc displacement and condylar alterations in this population reflects a structural and functional vulnerability of the temporomandibular region following oncologic disease and its treatment. The presence of soft-tissue and osseous changes show that the joint undergoes both inflammatory and degenerative remodeling, which may precede clinically evident dysfunction such as trismus or masticatory limitation. The need for proactive TMJ screening and multidisciplinary monitoring in oncologic care

is necessary. Since both CT and MRI are routinely acquired for staging, treatment planning, and follow-up of HNC, including a structured TMJ assessment within these scans would add minimal additional time yet provide valuable diagnostic information. Detecting early morphological changes may enable timely interventions (such as guided physiotherapy, jaw-opening exercises, and preventive splint therapy) to preserve mandibular mobility and prevent the onset of chronic pain or functional impairment.

Moreover, TMJ changes in cancer patients are not found to be isolated findings but rather part of a broader cascade involving muscle atrophy, fibrosis, and vascular compromise. Recognizing TMJ pathology early may therefore help clinicians anticipate complications affecting oral nutrition, speech, and hygiene. Combining TMJ evaluation into standard radiologic protocols and follow-up routines can transform incidental imaging findings into actionable information.

Several limitations should be considered in this study. First, disc displacement was recorded as present or absent because only closed-mouth MRI was available; grading displacement severity would require open-mouth dynamic imaging. The imaging variety in this patient population was inherently limited, as the study was retrospective and most patients had undergone imaging primarily for oncologic assessment rather than dedicated TMJ evaluation. Second, the absence of open-mouth MRI sequences represents a limitation, as it precluded differentiation between disc displacement with and without reduction. Recent studies have shown that static, closed-mouth MRI provides high diagnostic reliability for identifying disc displacement and joint alterations, and that dynamic imaging cannot fully replace static sequences but serves as a complementary tool.^{21,22}

Limitations

This retrospective study lacked direct clinical correlation parameters such as maximal mouth opening, pain intensity, and trismus degree, which limits functional interpretation of imaging findings. The head and neck cancer cohort was heterogeneous, including various tumor sites, histologic subtypes, and treatment modalities, which may have influenced the TMJ alterations observed. Furthermore, radiotherapy dosimetric data, such as dose distribution and target field configuration, were unavailable, precluding analysis of potential dose-response relationships. All imaging evaluations were performed by a single experienced oral and maxillofacial radiologist, and the assessment of intra- or interobserver agreement was not present. Future studies including multiple blinded observers would strengthen the reproducibility of these findings.

CONCLUSION

HNC is associated with TMJ disorders; patients with HNC showed more frequent disc displacement and a shift toward erosive condylar change. These findings are consistent with treatment- and disease-related vulnerability of joint tissues. Systematic TMJ evaluation within HNC care offers a feasible avenue to identify at-risk patients and to deploy timely, multidisciplinary interventions that preserve oral function and quality of life.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Kocaeli Health and Technology University Non-interventional Researches Ethics Committee (Date: 10.03.2025, Decision No: 2025-128).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Acknowledgements

This study is dedicated to the memory of Mehmet Uğur Bozkurt, whose courage during his battle with cancer and its many complications remains a lasting source of strength.

REFERENCES

- Mesia R, Iglesias L, Lambea J, et al. SEOM clinical guidelines for the treatment of head and neck cancer (2020). *Clin Transl Oncol*. 2021;23(5):913-921. doi:10.1007/s12094-020-02533-1
- Alfouzan AF. Radiation therapy in head and neck cancer. *Saudi Med J*. 2021;42(3):247-254. doi:10.15537/smj.2021.42.3.20210660
- Alterio D, Marvaso G, Ferrari A, Volpe S, Orecchia R, Jereczek-Fossa BA. Modern radiotherapy for head and neck cancer. *Semin Oncol*. 2019;46(3):233-245. doi:10.1053/j.seminoncol.2019.07.002
- Guigay J, Le Caer H, Ortholan C, Aupérin A, Michel C, Mertens C. Treatment of inoperable elderly head and neck cancer patients. *Curr Opin Oncol*. 2019;31(3):152-159. doi:10.1097/CCO.0000000000000526
- Wierzbicka M, Bartochowska A, Strnad V, et al. The role of brachytherapy in the treatment of squamous cell carcinoma of the head and neck. *Eur Arch Otorhinolaryngol*. 2016;273(2):269-276. doi:10.1007/s00405-014-3332-8
- Yuwanati M, Gondivkar S, Sarode SC, et al. Oral health-related quality of life in oral cancer patients: systematic review and meta-analysis. *Future Oncol*. 2021;17(8):979-990. doi:10.2217/fon-2020-0881
- Kalita O, Hrabalek L, Halaj M, et al. Very late complications of oncotherapy in glioblastoma patients: a case series. *Biomed Pap Med Fac Univ Palacky Olomouc Czech Repub*. 2022;166(2):236-241. doi:10.5507/bp.2021.012
- Borrelli MR, Shen AH, Lee GK, Momeni A, Longaker MT, Wan DC. Radiation-induced skin fibrosis: pathogenesis, current treatment options, and emerging therapeutics. *Ann Plast Surg*. 2019;83(4S Suppl 1):S59-S64. doi:10.1097/SAP.0000000000002098
- Gupta A, Peyrin-Biroulet L, Ananthakrishnan AN. Risk of cancer recurrence in patients with immune-mediated diseases with use of immunosuppressive therapies: an updated systematic review and meta-analysis. *Clin Gastroenterol Hepatol*. 2024;22(3):499-512.e6. doi:10.1016/j.cgh.2023.07.027
- Romero-Reyes M, Bassiur JP. Temporomandibular disorders, bruxism and headaches. *Neurol Clin*. 2024;42(2):573-584. doi:10.1016/j.ncl.2023.12.010

11. Garstka AA, Kozowska L, Kijak K, et al. Accurate diagnosis and treatment of painful temporomandibular disorders: a literature review supplemented by own clinical experience. *Pain Res Manag.* 2023;2023:1002235. doi:10.1155/2023/1002235
12. Singer SR, Mupparapu M. Temporomandibular joint imaging. *Dent Clin North Am.* 2023;67(2):227-241. doi:10.1016/j.cden.2022.11.001
13. Omami G, Miller CS. Imaging evaluation of the temporomandibular joint. *Dent Clin North Am.* 2024;68(2):357-373. doi:10.1016/j.cden.2023.10.001
14. Office of the Surgeon General (US). Bone Health and Osteoporosis: A Report of the Surgeon General. Rockville (MD): Office of the Surgeon General (US); 2004. Chapter 2, The Basics of Bone in Health and Disease. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK45504/>
15. Hashida N, Shamoto H, Maeda K, Wakabayashi H. Impact of geniohyoid and masseter muscle masses on dysphagia after salvage surgery and radiotherapy in head and neck cancer. *Sci Rep.* 2021;11(1):2278. doi:10.1038/s41598-021-82039-0
16. Somay E, Topkan E, Pehlivan UA, et al. The use of pre-chemoradiotherapy total masseter muscle volume as a novel predictor of radiation-induced trismus in locally advanced nasopharyngeal carcinoma patients. *Tomography.* 2024;10(1):79-89. doi:10.3390/tomography10010007
17. Wallace JD, Calvo RY, Lewis PR, et al. Sarcopenia as a predictor of mortality in elderly blunt trauma patients: comparing the masseter to the psoas using computed tomography. *J Trauma Acute Care Surg.* 2017;82(1):65-72. doi:10.1097/TA.0000000000001297
18. Vlăduțu D, Popescu SM, Mercuț R, et al. Associations between bruxism, stress, and manifestations of temporomandibular disorder in young students. *Int J Environ Res Public Health.* 2022;19(9):5415. doi:10.3390/ijerph19095415
19. Pauli N, Fagerberg Mohlin B, Mejersjö C, Finizia C. Temporomandibular disorder as risk factor for radiation-induced trismus in patients with head and neck cancer. *Clin Exp Dent Res.* 2022;8(1):123-129. doi:10.1002/cre2.511
20. Klein RE. Onchologic stress: its causality in myofascial pain disorder or in temporomandibular disorder. *Am J Hosp Palliat Care.* 1998;15(2):117-121. doi:10.1177/104990919801500211
21. Tresoldi M, Dias R, Bracci A, Segù M, Guarda-Nardini L, Manfredini D. Magnetic resonance imaging evaluation of closed-mouth TMJ disc-condyle relationship in a population of patients seeking for temporomandibular disorders advice. *Pain Res Manag.* 2021;2021:5565747. doi:10.1155/2021/5565747
22. Kamel ZSAS, El-Shafey MHR, Hassanien OA, et al. Can dynamic magnetic resonance imaging replace static magnetic resonance sequences in evaluation of temporomandibular joint dysfunction? *Egypt J Radiol Nucl Med.* 2021;52(1):19. doi:10.1186/s43055-020-00396-8

Relationship between the success of arthrocentesis and preoperative magnetic resonance imaging findings

 Ezgi Yüceer Çetiner¹,  Dilara Kazan¹,  Burcu Baş Akkor²

¹Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Bahçeşehir University, İstanbul, Türkiye

²Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Ondokuz Mayıs University, İstanbul, Türkiye

Cite this article as: Yüceer Çetiner E, Kazan D, Baş Akkor B. Relationship between the success of arthrocentesis and preoperative magnetic resonance imaging findings. *J Health Sci Med.* 2025;8(6):1149-1154.

Received: 08.09.2025

Accepted: 15.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to investigate the relationship between preoperative magnetic resonance imaging (MRI) findings and clinical outcomes following arthrocentesis in patients with temporomandibular joint disorders.

Methods: A total of 67 patients who had been diagnosed with disc position abnormalities and degenerative temporomandibular joint conditions based on the Diagnostic Criteria for Temporomandibular Disorders (DC/TMD), had preoperative MRI scans, and were treated with arthrocentesis were included in the study. The recorded preoperative variables included age, sex, pain, and maximum mouth opening (MMO). Degree of disc displacement and joint effusion were assessed using preoperative magnetic resonance images. Bone changes were evaluated using both preoperative magnetic resonance and panoramic radiography images. Arthrocentesis was considered successful when the MMO exceeded 35 mm and the pain score was less than 3.

Results: At the 1-week and 3-month follow-ups, clinical examinations showed a significant reduction in temporomandibular joint (TMJ) pain and an increase in MMO ($p < 0.001$). The total success rate increased from 19.4% at the 1-week follow-up to 44.8% at 3 months, indicating a significant improvement over time. MRI findings, including disc displacement, effusion, and bone changes, showed no statistically significant association with treatment success. Although patients without effusion showed lower VAS scores at 1 week, this difference did not persist at 3 months.

Conclusion: Arthrocentesis is a successful intervention for treating TMD, resulting in significant reductions in pain and functional limitation. However, commonly evaluated preoperative MRI features do not appear to predict clinical success.

Keywords: Arthrocentesis, magnetic resonance imaging, temporomandibular disorder

INTRODUCTION

Temporomandibular disorder (TMD) refers to a chronic painful musculoskeletal condition affecting approximately 28% to 47% of the adult population.^{1,2} Common clinical findings include difficulty in mandibular movements, pain, limited jaw mobility, and clicking.³ Disc displacements (with or without reduction) and degenerative joint diseases are the two most frequent subgroups of patients with TMDs.⁴ Although clinical evaluation remains the cornerstone of TMD diagnosis, MRI provides valuable anatomical insights, enabling clinicians to identify soft tissue abnormalities like disc position, fluid accumulation, and structural changes in the condyle with high accuracy.^{5,6}

Initial management of TMD primarily focuses on conservative approaches aimed at reducing pain, improving joint function, and preventing disease progression. These include a soft diet, behavioral modification to reduce parafunctional habits, pharmacological therapy (e.g., nonsteroidal anti-inflammatory drugs and muscle relaxants), occlusal splint therapy, physical therapy modalities such as jaw exercises,

massage, thermotherapy, and low-level laser therapy.^{7,8} When first-line treatment approaches fail to provide adequate symptom relief, arthrocentesis is considered a viable option. Arthrocentesis, first introduced in 1991, has become widely accepted as a conservative and beneficial method for managing temporomandibular disorders. Arthrocentesis involves lavage of the upper joint space to eliminate inflammatory mediators, reduce intra-articular pressure, and restore mandibular mobility. The procedure is known to be especially beneficial in the early stages of internal derangement, where it may break down adhesions, alleviate pain, and improve joint function.⁹⁻¹²

Although arthrocentesis has shown high clinical success rates, treatment outcomes can vary depending on individual clinical and radiological factors. MRI findings may offer prognostic value by revealing specific joint features that influence response to intervention. Therefore, understanding the relationship between preoperative MRI characteristics and clinical success is essential to optimizing patient selection and treatment outcomes.¹³

Corresponding Author: Ezgi Yüceer Çetiner, ezgi.yuceercetiner@bau.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

This study aimed to evaluate the clinical success of arthrocentesis in patients with temporomandibular joint disorders, based on pain reduction and improvement in mouth opening, and to determine whether specific preoperative MRI findings are associated with treatment outcomes.

METHODS

Ethics

The retrospective data from patients diagnosed with disc displacement with/without reduction and degenerative joint disease according to diagnostic criteria for TMD (DC/TMD), who underwent arthrocentesis and had preoperative MRI scans, were included in this study. All treatments were performed at the Department of Oral and Maxillofacial Surgery, Faculty of Dentistry, Ondokuz Mayıs University between 2016 and 2018. The study protocol was approved by the Ondokuz Mayıs University Clinical Researches Ethics Committee (Date: 15.03.2019, Decision No: KA EK2019/259). All patients provided written informed consent. All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Data Collection

All patients underwent clinical evaluation following DC/TMD.⁴ The inclusion criteria were patients aged 16 years or older diagnosed with intra-articular temporomandibular joint disorders; the presence of joint-related pain and/or limitation of mandibular movement unresponsive to conservative treatments for at least 3 months, availability of preoperative MRI of the temporomandibular joint, and patients who underwent arthrocentesis as the first-line surgical intervention. Patients were excluded if they had muscle-origin pain, prior jaw trauma, or TMJ surgery, and systemic conditions known to impact joint function. Demographic information, including age and sex, was recorded for each patient. Clinical parameters were pain intensity and maximum mouth opening (MMO). Pain was recorded using a Visual Analog Scale (VAS; 0-10). MMO was measured with a ruler between the incisal edges of the maxillary and mandibular incisors.¹⁴ The follow-up evaluations were conducted at the 1st week and the 3rd month. Treatment effectiveness was defined by achieving a maximum mouth opening exceeding 35 mm, coupled with a pain score below 3 on the VAS.¹⁵

Joint effusion was identified on MRI by the presence of a bright hyperintense area observed inside the joint spaces. The degree of effusion was classified according to predefined grading criteria: absent (no visible fluid in the joint space); mild (a thin or linear area observed along the joint surfaces indicating a minimal amount of fluid; [Figure 1](#)); moderate (focal areas representing localized collections of fluid, without complete filling of the joint compartments; [Figure 2](#)); severe (diffuse and abundant high-signal intensity occupying both joint spaces, often extending into the retrodiscal area and potentially distending the joint capsule; [Figure 3](#)).¹⁶

The degree of disc displacement was classified into three groups based on MRI findings: no displacement, displacement with reduction, and displacement without reduction. Bone changes were recorded from both preoperative magnetic

resonance and panoramic radiography images and classified into three categories: no bone change, cortical erosion, and cortico-cancellous erosion.

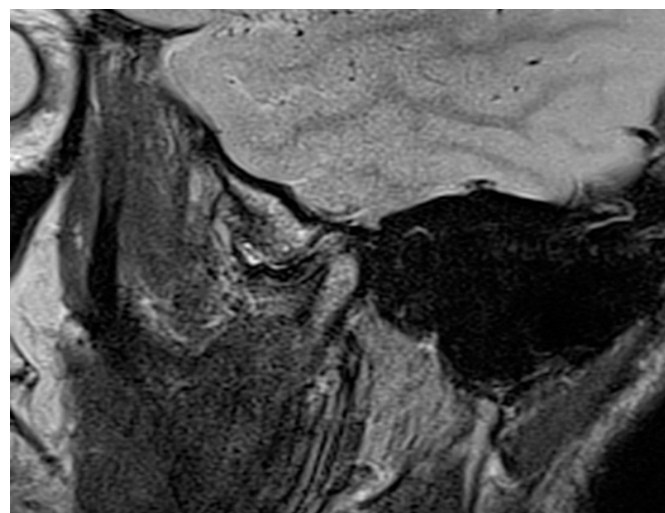


Figure 1. Sagittal T2-weighted magnetic resonance views showing mild effusion in the temporomandibular joint

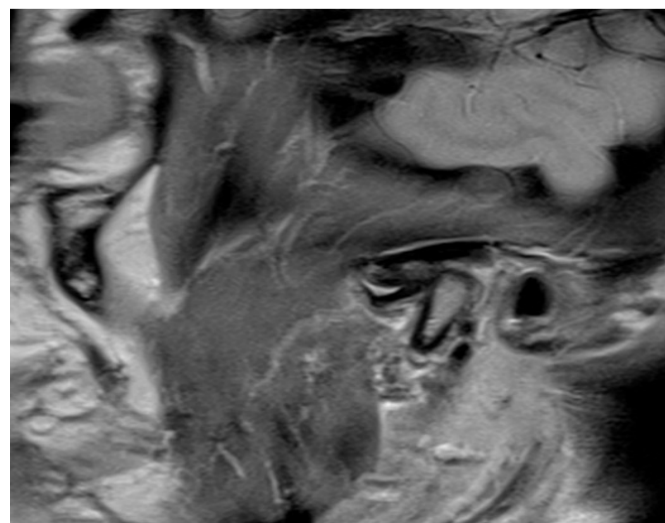


Figure 2. Sagittal T2-weighted magnetic resonance views showing moderate effusion in the temporomandibular joint

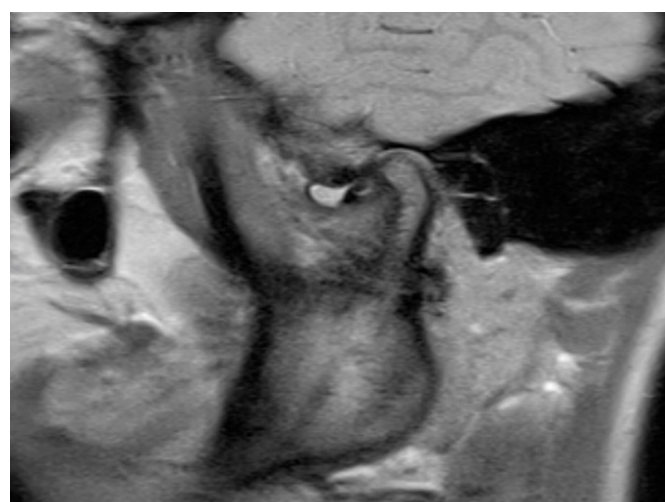


Figure 3. Sagittal T2-weighted magnetic resonance views showing severe effusion in the temporomandibular joint

Magnetic Resonance Imaging

0.5 Tesla scanner (SIGNA, GE, Milwaukee, WI, USA) with a 6×8 cm surface coil, 3-mm slice thickness, a 15-cm field of view, and matrix sizes of 256×192 or 256×256 was used for magnetic resonance imaging. T2-weighted images were acquired with TR/TE of 1500/20 ms or 1500/80 ms, and T1-weighted images with TR/TE between 200/11 ms and 340/17 ms. Oblique sagittal and coronal sections were obtained to assess the disc-condyle relationship. Bilateral TMJ images were captured in both resting and fully opened positions.

Arthrocentesis Procedure

All patients received a single-session arthrocentesis under sedation, following an auriculotemporal nerve block with 2 ml of local anesthetic administered to the preauricular area and upper joint space. Arthrocentesis was performed according to the classical two-needle technique described by Nitzan et al.¹² Needle insertion points were determined using a reference line from the tragus to the lateral canthus. The first puncture site was located approximately 10 mm anterior and 2 mm inferior to the tragus along this reference line, while the second was positioned about 20 mm anterior and 8 mm inferior to the same line. A 21-gauge needle was placed posteriorly to inject Ringer's lactate into the upper joint space, followed by a 20-gauge needle anteriorly for fluid outflow. After the procedure, a soft diet was recommended. Tenoxicam 20 mg tablets (Deva Holding Inc., İstanbul, Türkiye) were administered twice daily for 10 days postoperatively to control pain and reduce inflammation. Stabilization splints were prescribed for all patients and used for a minimum of one month following the intervention.

Statistical Analysis

All data evaluations were performed using IBM SPSS Statistics software (version 23.0; IBM Corp., Armonk, NY, USA). The Kolmogorov-Smirnov test was used to evaluate normal distribution. Depending on data distribution and group characteristics, different statistical methods were applied. Qualitative variables were analyzed using Pearson's Chi-Square test or Fisher's exact test with Monte Carlo adjustment where appropriate. For comparisons involving repeated measurements of non-normally distributed numerical data, the Friedman test was employed. If a significant difference was detected, post-hoc analysis was conducted with Bonferroni correction. Multiple group comparisons of variables with normal distribution were performed using One-way ANOVA, whereas the Kruskal-Wallis H test was applied for non-parametric data. The McNemar test was utilized to compare changes in qualitative variables over time. Correlations were evaluated using Spearman's rho coefficient. Quantitative variables were presented as mean±standard deviation and median (minimum-maximum), whereas qualitative variables were reported as frequency and percentage (n, %). A p-value below 0.05 was considered indicative of statistical significance.

RESULTS

67 patients (60 females, 7 males) with temporomandibular joint disorders were evaluated in the study. Ages ranged from 16 to 76 years, with a mean value of 33.13±13.57. According

to preoperative MRI, 26.9% had no displacement, 31.3% displacement with reduction, and 41.8% displacement without reduction. Joint effusion was absent in 16.4% of patients, mild in 41.8%, moderate in 20.9%, and severe in 20.9%. Bone changes were absent in 43.3%, while cortical erosion was observed in 38.8%, and cortico-cancellous erosion in 17.9% (Table 1).

Table 1. Descriptive statistics of the MRI findings

	Frequency (n)	Percentage (%)
Disc displacement		
None	18	26.9
With reduction	21	31.3
Without reduction	28	41.8
Effusion		
None	11	16.4
Mild	28	41.8
Moderate	14	20.9
Severe	14	20.9
Bone changes		
None	29	43.3
Cortical erosion only	26	38.8
Cortico-cancellous erosion	12	17.9

Data are presented as frequency (n) and percentage (%), MRI: Magnetic resonance imaging

There was a statistically significant improvement in both VAS and MMO scores between the preoperative period, the 1st week, and the 3rd month ($p<0.001$) (Table 2). No significant correlation was found between VAS and MMO scores at any time point ($p>0.05$). Additionally, comparisons of VAS and MMO scores across displacement degree, effusion, and bone change subgroups revealed no significant differences. Patients without effusion had significantly lower VAS scores at the 1st week compared to those with mild effusion ($p=0.028$).

Table 2. Comparison of VAS and MMO values before and after arthrocentesis

	Mean±SD	Median (min-max)	Test statistic	p
VAS				
Preoperative	7.27±1.48	8 (4-10) ^a	79.508	<0.001
1 st week	4.13±2.64	4 (0-10) ^b		
3 rd month	2.45±2.2	2 (0-8) ^c		
MMO				
Preoperative	28.39±6.66	27 (15-43) ^c	68.420	<0.001
1 st week	31.96±6.58	32 (16-50) ^b		
3 rd month	34.84±6.46	35 (25-50) ^a		

Mean±SD; Median (minimum-maximum), the Friedman test was used. Different superscript letters (a, b, c) indicate statistically significant differences between time points ($p<0.05$). VAS: Visual Analog Scale, MMO: Maximum mouth opening, SD: Standard deviation, min: Minimum, max: Maximum

When pain reduction and MMO increase were evaluated separately, the success rates were 40.3% and 32.8%, respectively, at the 1st week, and 65.7% and 59.7% at the 3rd month. However, when both criteria were considered together, the total success rate was calculated as 19.4% at the 1st week and 44.8% at the

3rd month. This difference in total success between the 1st week and the 3rd month was found to be statistically significant ($p=0.001$) (Table 3). Subgroup analysis showed no statistically significant relationship between success rate and the degree of disc displacement, effusion, or bone changes at either the 1st week or the 3rd month ($p>0.05$ for all).

Table 3. Comparison of total success rates over time			
Time point	Failure (n, %)	Success (n, %)	p
1 st week	54 (80.6)	13 (19.4)	0.001 ^x
3 rd month	37 (55.2)	30 (44.8)	
McNemar test was used			

DISCUSSION

Temporomandibular disorders are multifactorial conditions that can lead to significant pain and dysfunction in the TMJ.¹⁷ Arthrocentesis has emerged as a minimally invasive and effective treatment modality, particularly for patients with intra-articular disorders such as disc displacement and degenerative joint disease who are unresponsive to conservative therapy.¹⁸ The present study evaluated the clinical effectiveness of arthrocentesis in patients with TMD and investigated whether specific preoperative MRI findings—such as disc displacement, effusion, or bone changes—could predict treatment success.

The findings of this study demonstrated a significant improvement in both pain levels and MMO after arthrocentesis, consistent with previous literature showing the efficacy of this technique in managing TMD symptoms.¹⁹⁻²³ We adopted the success criteria proposed by Baş et al.,¹⁵ which define clinical success as achieving a maximum mouth opening greater than 35 mm and a visual analog scale score lower than 3 at the follow-up evaluation. This combined criterion allows for a more comprehensive assessment of both functional improvement and pain relief, rather than evaluating each parameter in isolation. By applying these criteria, we aimed to ensure comparability with previous literature and to provide a clinically meaningful definition of treatment success from a patient-centered perspective. Similar to previous findings in the literature, our results demonstrated a considerable difference between the independent and combined success rates.^{16,24}

Despite the overall improvement, subgroup analyses revealed no statistically significant associations between treatment success and preoperative MRI findings, including degree of disc displacement, joint effusion, or osseous changes. Although patients without effusion showed significantly lower VAS scores at the 1st week compared to those with mild effusion, this difference did not persist at the 3rd month, nor did it influence combined success rates. These results suggest that while MRI remains a valuable tool for diagnosing intra-articular pathology, its prognostic utility in predicting the clinical success of arthrocentesis may be limited.

Several previous studies have questioned the prognostic value of MRI findings in predicting the clinical outcome of arthrocentesis.²⁵⁻²⁷ For example, Sato et al.²⁸ evaluated pre-

and post-treatment magnetic resonance images in patients undergoing arthroscopic lavage and reported no significant association between joint effusion and clinical improvement following the procedure. Similarly, Ekberg et al.¹³ reported no significant association between MRI-based disc position and clinical outcome, aligning with our findings that preoperative MRI features such as the degree of disc displacement, effusion, or bone changes were not predictive of treatment success following arthrocentesis. Thus, reliance on preoperative MRI findings for outcome prediction may not be justified.

Although certain studies, such as those by Emshoff and Rudisch²⁹ and Honda et al.,³⁰ have suggested that extensive joint effusion or condylar erosion may negatively affect treatment response, our data did not support these observations. These discrepancies may stem from variations in study design, imaging interpretation protocols, or follow-up duration. Furthermore, the lack of association between MRI findings and treatment outcomes in our study supports the notion that MRI, despite its diagnostic utility, may not be necessary for predicting arthrocentesis success. Several explanations may account for the lack of correlation between imaging findings and clinical success. MRI provides only a static representation of joint morphology and may not fully capture the dynamic pathophysiology of TMJ disorders. Moreover, pain and functional limitations in TMD are influenced by a complex interplay of biomechanical, psychological, and systemic factors that may not be evident on imaging alone.

Limitations

This study offers valuable insights through the use of standardized diagnostic criteria (DC/TMD and MRI), clearly defined inclusion parameters, and a dual outcome assessment combining pain relief and functional improvement. Nevertheless, there are a few limitations to this study. First, its retrospective design may introduce selection bias and limit the ability to draw causal inferences. Second, MRI interpretation was based on qualitative visual grading without interobserver reliability analysis or quantitative assessment, which may affect diagnostic precision. Additionally, patients could not be classified according to the Wilkes staging system because detailed clinical data required for accurate staging were not consistently available in the retrospective records. Further studies should adopt prospective designs with larger cohorts and employ standardized, preferably quantitative, MRI assessment protocols. Incorporating advanced imaging techniques such as dynamic MRI or cone-beam computed tomography (CBCT) could provide deeper insight into joint mechanics and enhance the ability to predict treatment response. Extended follow-up periods will also be necessary to assess the long-term stability and effectiveness of arthrocentesis.

CONCLUSION

As a result, arthrocentesis appears to be a clinically useful treatment option for patients with TMD, offering significant improvements in pain and jaw function over time. Notably, although MRI findings such as disc displacement, effusion, and bone erosion were prevalent, they did not significantly correlate with early or late treatment success. This suggests that

while imaging findings provide valuable diagnostic insight, they may not serve as reliable predictors of clinical outcomes. The significant reduction in pain and functional limitation even among patients with advanced displacement or effusion further underscores the potential effectiveness of the applied treatment approach-possibly indicating that therapeutic response is not solely dependent on the severity of structural changes observed in MRI. The absence of correlation between VAS and MMO scores also supports the multifactorial nature of TMJ disorders, where subjective pain experience and mechanical limitation may evolve independently.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study received the necessary ethical approval from the Ondokuz Mayıs University Clinical Researches Ethics Committee (Date: 15.03.2019, Decision No: KAEK2019/259).

Informed Consent

This study was conducted retrospectively using existing patient data. Therefore, additional written informed consent was not required. However, all patients had provided written consent for the use of their clinical records for research purposes at the time of treatment.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Guo C, Shi Z, Revington P. Arthrocentesis and lavage for treating temporomandibular joint disorders. *Cochrane Database Syst Rev*. 2009; (4):CD004973. doi:10.1002/14651858.CD004973.pub2
- Jagur O, Kull M, Leibur E, et al. The associations of TMJ pain and bone characteristics on the activities of daily living. *Open J Stomatol*. 2012; 2(4):237-243.
- Paesani D, Westesson PL, Hatala M, Tallents RH, Kurita K. Prevalence of temporomandibular joint internal derangement in patients with craniomandibular disorders. *Am J Orthod Dentofacial Orthop*. 1992; 101(1):41-47. doi:10.1016/0889-5406(92)70080-T
- Schiffman E, Ohrbach R, Truelove E, et al. Diagnostic criteria for temporomandibular disorders (DC/TMD) for clinical and research applications: recommendations of the international RDC/TMD consortium network* and orofacial pain special interest group. *J Oral Facial Pain Headache*. 2014;28(1):6-27. doi:10.11607/jop.1151
- Ahmad M, Hollender L, Anderson Q, et al. Research diagnostic criteria for temporomandibular disorders (RDC/TMD): development of image analysis criteria and examiner reliability for image analysis. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2009;107(6):844-860. doi:10.1016/j.tripleo.2009.02.023
- Sano T, Westesson PL. Magnetic resonance imaging of the temporomandibular joint. Increased T2 signal in the retrodiskal tissue of painful joints. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 1995;79(4):511-516. doi:10.1016/s1079-2104(05)80138-6
- Hosgor H, Bas B, Celenk C. A comparison of the outcomes of four minimally invasive treatment methods for anterior disc displacement of the temporomandibular joint. *Int J Oral Maxillofac Surg*. 2017;46(11):1403-1410. doi:10.1016/j.ijom.2017.05.010
- Wieckiewicz M, Boening K, Wiland P, Shiao YY, Paradowska-Stolarz A. Reported concepts for the treatment modalities and pain management of temporomandibular disorders. *J Headache Pain*. 2015;16:106. doi:10.1186/s10194-015-0586-5
- Milam SB, Schmitz JP. Molecular biology of temporomandibular joint disorders: proposed mechanisms of disease. *J Oral Maxillofac Surg*. 1995;53(12):1448-1454. doi:10.1016/0278-2391(95)90675-4
- Zardeneta G, Milam SB, Lee T, Schmitz JP. Detection and preliminary characterization of matrix metalloproteinase activity in temporomandibular joint lavage fluid. *Int J Oral Maxillofac Surg*. 1998; 27(5):397-403. doi:10.1016/s0901-5027(98)80072-6
- Reston JT, Turkelson CM. Meta-analysis of surgical treatments for temporomandibular articular disorders: a reply to the discussants. *J Oral Maxillofac Surg*. 2003;61(6):737-738. doi:10.1053/joms.2003.50169
- Nitzan DW, Dolwick MF, Martinez GA. Temporomandibular joint arthrocentesis: a simplified treatment for severe, limited mouth opening. *J Oral Maxillofac Surg*. 1991;49(11):1163-1170. doi:10.1016/0278-2391(91)90409-f
- Ekberg E, Hansson LG, List T, Eriksson L, Sahlström LE, Petersson A. Can MRI observations predict treatment outcome of lavage in patients with painful TMJ disc displacement without reduction? *J Oral Maxillofac Res*. 2015;6(1):e5. doi:10.5037/jomr.2014.6105
- Güler N, Uçkan S, Imirzalioglu P, Açikgözoğlu S. Temporomandibular joint internal derangement: relationship between joint pain and MR grading of effusion and total protein concentration in the joint fluid. *Dentomaxillofac Radiol*. 2005;34(3):175-181. doi:10.1259/dmfr/49181266
- Bas B, Yuceer E, Kazan D, Gurbanov V, Kutuk N. Clinical and intra-operative factors affecting the outcome of arthrocentesis in disc displacement without reduction: a retrospective study. *J Oral Rehabil*. 2019;46(8):699-703. doi:10.1111/joor.12808
- Larheim TA, Westesson PL, Sano T. MR grading of temporomandibular joint fluid: association with disk displacement categories, condyle marrow abnormalities and pain. *Int J Oral Maxillofac Surg*. 2001;30(2):104-112. doi:10.1054/ijom.2000.0017
- Nitzan DW, Naaman HL. Arthrocentesis: what, when, and why? *Atlas Oral Maxillofac Surg Clin North Am*. 2022;30(2):137-145. doi:10.1016/j.cxom.2022.06.008
- Hu Y, Liu S, Fang F. Arthrocentesis vs conservative therapy for the management of TMJ disorders: a systematic review and meta-analysis. *J Stomatol Oral Maxillofac Surg*. 2023;124(1S):101283. doi:10.1016/j.jomas.2022.09.004
- Alpaslan C, Dolwick MF, Heft MW. Five-year retrospective evaluation of temporomandibular joint arthrocentesis. *Int J Oral Maxillofac Surg*. 2003;32(3):263-267. doi:10.1054/ijom.2003.0371
- Goudot P, Jaquinet AR, Hugonnet S, Haefliger W, Richter M. Improvement of pain and function after arthroscopy and arthrocentesis of the temporomandibular joint: a comparative study. *J Craniomaxillofac Surg*. 2000;28(1):39-43. doi:10.1054/jcms.1999.0103
- Nitzan DW, Price A. The use of arthrocentesis for the treatment of osteoarthritic temporomandibular joints. *J Oral Maxillofac Surg*. 2001; 59(10):1154-1160. doi:10.1053/joms.2001.26716
- Alpaslan GH, Alpaslan C. Efficacy of temporomandibular joint arthrocentesis with and without injection of sodium hyaluronate in treatment of internal derangements. *J Oral Maxillofac Surg*. 2001;59(6):613-619. doi:10.1053/joms.2001.23368
- Emshoff R. Clinical factors affecting the outcome of arthrocentesis and hydraulic distension of the temporomandibular joint. *Oral Surg Oral Med Oral Pathol Oral Radiol Endod*. 2005;100(4):409-414. doi:10.1016/j.tripleo.2004.12.024
- Nishimura M, Segami N, Kaneyama K, Suzuki T. Prognostic factors in arthrocentesis of the temporomandibular joint: evaluation of 100 patients with internal derangement. *J Oral Maxillofac Surg*. 2001;59(8):874-878. doi:10.1053/joms.2001.25019
- Emshoff R, Rudisch A. Are internal derangement and osteoarthritis linked to changes in clinical outcome measures of arthrocentesis of the temporomandibular joint? *J Oral Maxillofac Surg*. 2003;61(10):1162-1170. doi:10.1016/s0278-2391(03)00676-1

26. Iwase H, Sasaki T, Asakura S, et al. Characterization of patients with disc displacement without reduction unresponsive to nonsurgical treatment: a preliminary study. *J Oral Maxillofac Surg.* 2005;63(8):1115-1122. doi:10.1016/j.joms.2005.04.005
27. Smolka W, Iizuka T. Arthroscopic lysis and lavage in different stages of internal derangement of the temporomandibular joint: correlation of preoperative staging to arthroscopic findings and treatment outcome. *J Oral Maxillofac Surg.* 2005;63(4):471-478. doi:10.1016/j.joms.2004.07.021
28. Sato J, Segami N, Kaneyama K, Nishimura M, Suzuki T. Importance of the changes in joint effusion shown by magnetic resonance imaging before and after arthroscopic lysis and lavage of the temporomandibular joint. *Br J Oral Maxillofac Surg.* 2002;40(1):72-75. doi:10.1054/bjom.2001.0737
29. Emshoff R, Rudisch A. Temporomandibular joint internal derangement and osteoarthritis: are effusion and bone marrow edema prognostic indicators for arthrocentesis and hydraulic distention? *J Oral Maxillofac Surg.* 2007;65(1):66-73. doi:10.1016/j.joms.2005.11.113
30. Honda K, Yasukawa Y, Fujiwara M, Abe T, Urade M. Causes of persistent joint pain after arthrocentesis of temporomandibular joint. *J Oral Maxillofac Surg.* 2011;69(9):2311-2315. doi:10.1016/j.joms.2011.02.025

Feasibility of androgen receptor pathway inhibitors in prostate cancer: impact of comorbidities, age, and concomitant medication use-a multicenter cohort study

¹Nurullah İlhan¹, ²Akif Doğan¹, ³Erdem Çubukçu², ⁴İlhan Hacıbekiroğlu³, ⁵Murat Araz⁴,
⁶Hacı Mehmet Türk⁵, ⁷Mehmet Beşiroğlu⁶, ⁸İlker Nihat Ökten⁶,
⁹Süleyman Baş¹, ¹⁰Alpaslan Tanoğlu⁷, ¹¹Mahmut Gümüş⁶

¹Department of Medical Oncology, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, İstanbul, Türkiye

²Division of Medical Oncology, Department of Internal Medicine, Faculty of Medicine, Bursa Uludağ University, Bursa, Türkiye

³Division of Medical Oncology, Department of Internal Medicine, Faculty of Medicine, Sakarya University, Sakarya, Türkiye

⁴Department of Medical Oncology, Department of Internal Medicine, Faculty of Medicine, Necmettin Erbakan University, Konya, Türkiye

⁵Department of Medical Oncology, Department of Internal Medicine, Faculty of Medicine, Bezmialem Vakıf University, İstanbul, Türkiye

⁶Department of Medical Oncology, İstanbul Medeniyet University, Göztepe Training and Research Hospital, İstanbul, Türkiye

⁷Department of Internal Medicine, Faculty of Medicine, Bahçeşehir University, İstanbul, Türkiye

Cite this article as: İlhan N, Doğan A, Çubukçu E, et al. Feasibility of androgen receptor pathway inhibitors in prostate cancer: impact of comorbidities, age, and concomitant medication use-a multicenter cohort study. *J Health Sci Med.* 2025;8(6):1155-1163.

Received: 02.10.2025

Accepted: 16.10.2025

Published: 25.10.2025

ABSTRACT

Aims: Androgen receptor pathway inhibitors (ARPIs) have improved outcomes in advanced prostate cancer. Still, evidence regarding their feasibility in older patients with comorbidities and multiple concomitant medications remains limited, as such populations are often underrepresented in clinical trials.

Methods: This multicenter, retrospective cohort study included 538 prostate cancer patients diagnosed between January 2021 and November 2023 across seven centers in Türkiye. Demographic, clinical, pathological, and treatment data were extracted from institutional records. Patients were stratified by treatment type [androgen deprivation therapy (ADT), androgen receptor pathway inhibitor (ARPI), chemotherapy (CT)] and by the number of medications used ($\geq$ three vs. <3). The primary outcome was overall survival (OS); secondary outcomes included follow-up duration, tumor grade, and metastatic distribution.

Results: The mean age at diagnosis was 70.3 years, and 59.5% of patients had comorbidities. Concomitant medication use of three or more drugs, was observed in 34.4%. Metastatic disease was present in 82.3% of cases, most commonly involving the bone (62.8%). ARPI therapy was administered to 72.7% of patients, ADT alone to 18.6%, and CT to 8.7%. Patients receiving ARPIs had higher comorbidity and concomitant medication rates and more frequent metastases than those receiving ADT. Still, they achieved significantly longer median follow-up (901 vs. 470 days, $p<0.001$) and prolonged OS. Patients with concomitant medication use of three or more drugs also showed a longer follow-up (1081 vs. 573 days, $p<0.001$), lower mortality (32.4% vs. 47.3%, $p=0.001$), and a higher proportion of grade 1 tumors compared with patients using fewer than three medications.

Conclusion: In this large real-world cohort, ARPIs were found to be feasible and effective, even in elderly, comorbid, and medication use of three or more patients, supporting their use beyond traditional trial populations.

Keywords: Prostate cancer, ARPI, comorbidity, medication, real-world study

INTRODUCTION

Prostate cancer remains one of the most common malignancies among men worldwide, with incidence and mortality increasing with advancing age, mainly due to global demographic shifts. Older patients (≥ 65 years) represent a large proportion of prostate cancer diagnoses, and age is a well-established prognostic factor influencing both disease presentation and treatment outcomes.^{1,2} As life expectancy increases, so too does the burden of comorbidities and the use of multiple concurrent medications in this population.^{3,4}

The use of multiple concurrent medications is highly prevalent among elderly cancer patients, with prevalence estimates varying depending on the definition and population studied,

but frequently exceeding 50%.^{5,6} In prostate cancer specifically, comorbid conditions such as cardiovascular disease, diabetes, and chronic kidney disease are common and are known to affect survival independently of cancer-specific factors adversely.^{7,8}

Novel androgen receptor pathway inhibitors (ARPIs), which target androgen receptor signaling, have significantly advanced the treatment of advanced prostate cancer. Randomized controlled trials have demonstrated survival advantages over androgen deprivation therapy (ADT) alone, with manageable toxicity profiles, even in older patients.^{9,10} However, patients with high comorbidity burden and those

Corresponding Author: Nurullah İlhan, nurullahilhan07@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

receiving multiple medications are often underrepresented in clinical trials, underscoring the importance of real-world safety and effectiveness data in these subgroups for clinical decision-making.¹¹

This study aims to address this gap by evaluating the feasibility, safety, and effectiveness of ARPI therapy in older prostate cancer patients with high comorbidity burden and use of multiple concurrent medications. We compared these patients with those using fewer medications or receiving alternative therapies [ADT or chemotherapy (CT)] with respect to metastatic burden, tumor grade, treatment type, follow-up duration, and survival outcomes.

METHODS

Ethics

The study was carried out with the permission of the Şehit Prof. Dr. İlhan Varank Training and Research Hospital Ethics Committee (Date: 19.08.2024, Decision No: 2024/262). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Study Design and Population

This multicenter, retrospective observational study included 538 patients diagnosed with prostate cancer between January 2021 and November 2023 from 7 participating centers in Türkiye. Eligible patients were adults (≥ 18 years) with histopathologically confirmed prostate cancer and complete baseline data available in institutional records. Patients who initiated systemic therapy in the metastatic hormone-sensitive (mHSPC) setting were included. Patients with incomplete clinical information, missing follow-up data, or concurrent malignancies were excluded from the analysis.

Data Collection

This multicenter, retrospective observational study included 538 patients diagnosed with prostate cancer between January 2021 and November 2023 from 7 participating centers in Türkiye. Eligible patients were adults (≥ 18 years) with histopathologically confirmed prostate cancer and complete baseline data available in institutional records. Demographic variables (age at diagnosis, family history), clinical characteristics (comorbidity status, medication use), tumor-related parameters (histological grade, metastatic stage at diagnosis), and treatment-related data [ADT, novel androgen receptor pathway inhibitor (ARPI) therapy, CT, and local treatments such as surgery or radiotherapy] were obtained from electronic data systems and manually recorded patient files. Medication use was evaluated by the number of concomitant medications at baseline, and patients were stratified according to a threshold of ≥ 3 medications.

Outcomes and Follow-Up

The primary aim of the study was to evaluate the real-world feasibility and treatment patterns of ARPIs, ADT, and CT in older and comorbid patients with metastatic prostate cancer. Outcomes were analyzed descriptively to characterize patient distribution by age, comorbidity, tumor grade, medication use, and treatment type. Overall survival (OS), defined as the time from diagnosis to death from any cause, was assessed as

a secondary endpoint to provide contextual information and to explore how baseline factors such as comorbidity burden and the number of concomitant medications might influence patient outcomes. Mortality events were verified through hospital records and the national death registry, and the median follow-up duration was recorded for the entire cohort.

Statistical Analysis

All data analyses were performed using IBM SPSS Statistics for Windows, Version 25.0 (IBM Corp., Armonk, NY, USA). The normality of distribution for continuous variables was assessed using the Kolmogorov-Smirnov test. Continuous variables with normal distribution were expressed as mean $\pm$ standard deviation (SD), while those without normal distribution were presented as median [interquartile range (IQR)]. For comparisons between two independent groups, the independent t-test was used for normally distributed variables, and the Mann-Whitney U test was used for non-normally distributed variables. For comparisons among three or more groups, the ANOVA test was used for normally distributed variables, with Bonferroni correction applied for post-hoc analyses when significant differences were observed. For non-normally distributed continuous variables, the Kruskal-Wallis test was applied, followed by appropriate post-hoc tests (e.g., Mann-Whitney U test). Chi-square test and Fisher's exact test were used for categorical variables. Survival analyses were conducted using the Kaplan-Meier method, with comparisons between groups performed using the Log-Rank test. Kaplan-Meier curves were generated to illustrate OS stratified by treatment type and medication use. To evaluate the independent effects of key clinical factors on OS, a multivariable Cox proportional hazards regression analysis was performed. Covariates included age at diagnosis, comorbidities, tumor grade, presence of metastasis, treatment type, and medication use (defined as use of three or more medications). Hazard ratios (HRs) and 95% confidence intervals (CIs) were reported. A p-value of less than 0.05 was considered statistically significant in all analyses.

RESULTS

A total of 538 patients diagnosed with prostate cancer were included in this multicenter study, conducted across seven centers between January 2021 and November 2023. The mean age at diagnosis was 70.3 ± 10.9 years. A positive family history of prostate cancer was reported in 3.9% of patients. Comorbid conditions were observed in 59.5% of the cohort, the most common being hypertension (46.5%), followed by coronary artery disease (21.4%), diabetes mellitus (19.9%), chronic obstructive pulmonary disease (COPD, 3.2%), and chronic kidney disease (CKD, 1.3%). Regarding medication use, 34.4% of patients were taking three or more concomitant medications, while 65.6% were taking fewer than three ([Table 1](#)).

Histological grade and metastatic stage were available for the majority of patients. Before therapy, 50.2% had grade 1 tumors, 25.1% had grade 2, and 24.7% had grade 3. Local treatment was administered to 26% of patients, with 13.6% receiving radiotherapy and 12.5% undergoing surgery ([Table 1](#)).

Table 1. Demographic, clinical, and treatment characteristics of prostate cancer patients

Characteristic	Patients (n=538)
Age at diagnosis (years), mean±SD	70.26±10.89
Family history	21 (3.9%)
Comorbidities	320 (59.5%)
Hypertension	250 (46.5%)
Coronary artery disease	115 (21.4%)
Diabetes	107 (19.9%)
COPD	17 (3.2%)
Chronic kidney disease	7 (1.3%)
Medication use	
≥3 medications	185 (34.4%)
<3 medications	353 (65.6%)
Grade before therapy	
1	270 (50.2%)
2	135 (25.1%)
3	133 (24.7%)
Local treatment type	140 (26%)
Radiotherapy	73 (13.6%)
Operation	67 (12.5%)
Metastatic sites	443 (82.3%)
Bone metastasis	338 (62.8%)
Lung metastasis	57 (10.6%)
Liver metastasis	21 (3.9%)
Nonregional lymph nodes	180 (33.5%)
Treatment type	
Only ADT	100 (18.6%)
ARPI	391 (72.7%)
Only chemotherapy	47 (8.7%)
Follow-up period (days), median (IQR)	667 (243.50-1655.55)
Exitus (death)	227 (42.2%)
Time from diagnosis to exitus (days), median (IQR)	772 (191-1629)

SD: Standard deviation, COPD: Chronic obstructive pulmonary disease, ECOG: Eastern cooperative oncology group, IQR: Interquartile range, ADT: Androgen deprivation therapy, ARPI: Androgen receptor pathway inhibitors

Metastatic disease was identified in 82.3% of patients. The most common sites were bone (62.8%), followed by non-regional lymph nodes (33.5%), lung (10.6%), and liver (3.9%) (Table 1).

Regarding systemic treatment, 18.6% of patients received ADT alone, 72.7% received novel ARPI therapy in combination with ADT, and 8.7% received CT in combination with ADT (Table 1).

The median follow-up duration was 667 days (IQR, 243.5-1655.6). During follow-up, 227 patients (42.2%) died. The median time from diagnosis to death was 772 days (IQR, 191-1629) (Table 1).

When comparing patients who died during follow-up (n=227) with those who survived (n=311), no significant difference in mean age at diagnosis was observed (70.3±8.6 vs. 70.3±12.3

years, p=0.987). A positive family history of prostate cancer was slightly more common among deceased patients (5.7% vs. 2.6%), but this did not reach statistical significance (p=0.062). The overall prevalence of comorbidities was comparable between groups (62.1% vs. 57.6%, p=0.288), with no significant differences in hypertension or coronary artery disease rates. However, diabetes mellitus was significantly more frequent among deceased patients compared with survivors (24.2% vs. 16.7%, p=0.031). No significant differences in the prevalence of COPD or CKD were observed between groups (Table 2).

Table 2. Comparison of demographic, clinical, and treatment characteristics between patients with and without exitus

Characteristic	Exitus (n=227)	No exitus (n=311)	p
Age at diagnosis (years), mean±SD	70.27±8.61	70.26±12.31	0.987 ^a
Family history	13 (5.7%)	8 (2.6%)	0.062 ^b
Comorbidities	141 (62.1%)	179 (57.6%)	0.288 ^b
Hypertension	111 (48.9%)	139 (44.7%)	0.334 ^b
CAD	49 (21.6%)	66 (21.2%)	0.919 ^b
Diabetes	55 (24.2%)	52 (16.7%)	0.031 ^{b,*}
COPD	5 (2.2%)	12 (3.9%)	0.278 ^b
CKD	1 (0.4%)	6 (1.9%)	0.248 ^c
Medication use			0.001 ^{b,*}
≥3 medications	60 (26.4%)	125 (40.2%)	
<3 medications	167 (73.6%)	186 (59.8%)	
Grade before therapy			0.011 ^{b,*}
1	103 (45.4%)	167 (53.7%)	
2	53 (23.3%)	82 (26.4%)	
3	71 (31.3%) ^k	62 (19.9%)	
Local treatment type	60 (26.4%)	80 (25.7%)	0.853 ^b
Radiotherapy[†]	27 (45%)	46 (57.5%)	0.143 ^b
Operation[‡]	33 (55%)	34 (42.5%)	
Metastatic sites	202 (89%)	241 (77.5%)	0.001 ^{b,*}
Bone metastasis	157 (69.2%)	181 (58.2%)	0.009 ^{b,*}
Lung metastasis	24 (10.6%)	33 (10.6%)	0.989 ^b
Liver metastasis	8 (3.5%)	13 (4.2%)	0.689 ^b
Nonregional LN	67 (29.5%)	113 (36.3%)	0.098 ^b
Treatment type			<0.001 ^{b,*}
Only ADT	9 (4%)	91 (29.3%) ^k	
ADT with ARPI	184 (81.1%) ^l	207 (66.6%)	
ADT with chemotherapy	34 (15%) ^m	13 (4.2%)	
Follow-up period (days), median (IQR)	772 (191-1629)	620 (290-1732)	0.519 ^d

SD: Standard deviation, CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, CKD: Chronic kidney disease, LN: Lymph nodes, ADT: Androgen deprivation therapy, IQR: Interquartile range, ARPI: Androgen receptor pathway inhibitors
a: Independent T test; b: Chi-square test; c: Fisher's exact test; d: Mann-Whitney U test. k,l,m: Letters indicate statistically significant differences between groups. [†]: This symbol is used to indicate that the percentages represent the distribution of patients according to their local treatment type categories. *: Statistically significant (p<0.05)

Importantly, patients who died were significantly less likely to be using three or more medications compared with survivors (26.4% vs. 40.2%, p=0.001), suggesting a potential inverse association between the use of three or more drugs

and OS (Table 2). This observation may reflect closer medical monitoring and more comprehensive management of comorbidities among survivors, rather than a direct protective effect of higher medication use.

Regarding tumor grade before therapy, a significantly higher proportion of deceased patients had grade 3 disease compared with survivors (31.3% vs. 19.9%, $p=0.011$), whereas grade 1 tumors were more common among survivors (53.7% vs. 45.4%). Local treatment rates were similar between groups (26.4% vs. 25.7%, $p=0.853$). Among patients who underwent local treatment, radiotherapy was more frequently administered to survivors (57.5% vs. 45.0%), while surgery was more common among deceased patients (55.0% vs. 42.5%); however, these differences did not reach statistical significance (Table 2).

Metastatic disease at diagnosis was significantly more frequent among deceased patients compared with survivors (89.0% vs. 77.5%, $p=0.001$), with bone metastases being particularly more prevalent in the deceased group (69.2% vs. 58.2%, $p=0.009$). No significant differences in the frequency of lung, liver, or nonregional lymph node metastases were observed between groups (Table 2).

Treatment distribution differed significantly between groups ($p<0.001$). Deceased patients were more likely to have received ARPI therapy (81.1% vs. 66.6%) or CT (15.0% vs. 4.2%). In contrast, ADT was administered more frequently alone to survivors (29.3% vs. 4.0%). The median follow-up duration

was slightly longer among deceased patients (772 days; IQR, 191-1629) compared with survivors (620 days; IQR, 290-1732), although this difference did not reach statistical significance ($p=0.519$) (Table 2).

When patients were stratified by treatment type, significant differences emerged across several clinical and demographic variables. Patients receiving ADT alone were significantly older at diagnosis compared with those in the ARPI group (73.7 ± 15.7 vs. 69.4 ± 9.0 years, $p=0.002$). A positive family history of prostate cancer was most frequent in the CT group (12.8%) and least frequent in the ARPI group (2.6%, $p=0.013$). Although the overall prevalence of comorbidities did not differ significantly across treatment groups ($p=0.253$), patients in the ARPI group exhibited significantly higher rates of hypertension (52.4%), coronary artery disease (25.8%), and diabetes mellitus (23.5%) compared with the other groups (all $p<0.001$). Chronic kidney disease was most common in the ADT group (4.0%, $p=0.042$).

Regarding medication use, patients in the ARPI group had the highest proportion of patients using three or more medications (43.7%), which was significantly higher than in the ADT (14.0%) and CT (0%) groups ($p<0.001$). Notably, all patients in the CT group were taking fewer than three medications. This pattern may reflect differences in comorbidity burden, treatment intensity, or patient selection across groups (Table 3).

Table 3. Comparison of demographic, clinical, and treatment characteristics by treatment type in prostate cancer patients

Characteristic	Only ADT (n=100)	ARPI (n=391)	CT (n=47)	p
Age at diagnosis (years), mean $\pm$ SD	73.69 $\pm$ 15.73 ^k	69.41 $\pm$ 9.03 ^k	70.11 $\pm$ 11.38	0.002 ^{b,*}
Family history	5 (5%)	10 (2.6%)	6 (12.8%) ^k	0.013 ^{b,*}
Comorbidities	54 (54%)	241 (61.6%)	25 (53.2%)	0.253 ^b
Hypertension	28 (28%)	205 (52.4%) ^k	17 (36.2%)	<0.001 ^{b,*}
CAD	9 (9%)	101 (25.8%) ^k	5 (10.6%)	<0.001 ^{b,*}
Diabetes	10 (10%)	92 (23.5%) ^k	5 (10.6%)	0.001 ^{b,*}
COPD	5 (2%)	10 (2.6%)	2 (4.3%)	0.448 ^b
CKD	4 (4%) ^k	2 (0.5%)	1 (2.1%)	0.042 ^{b,*}
Medication use				<0.001 ^{b,*}
≥3 medications	14 (14%)	171 (43.7%) ^k	0 (0%)	
<3 medications	86 (86%)	220 (56.3%)	47 (100%) ^l	
Grade before therapy				<0.001 ^{b,*}
1	37 (37%)	225 (57.5%) ^k	8 (17%)	
2	27 (27%)	106 (27.1%)	2 (4.3%) ^l	
3	36 (36%)	60 (15.3%)	37 (78.7%) ^m	
Local treatment type	20 (20%)	112 (28.6%)	8 (17%)	0.063 ^b
Radiotherapy [†]	17 (85%) ^k	55 (49.1%)	1 (12.5%)	<0.001 ^{b,*}
Operation [†]	3 (15%)	57 (50.9%)	7 (87.5%) ^l	
Metastatic sites	32 (32%)	375 (95.9%) ^k	36 (76.6%)	<0.001 ^{b,*}
Bone metastasis	20 (20%)	293 (74.9%) ^k	25 (53.2%)	<0.001 ^{b,*}
Lung metastasis	2 (2%)	54 (13.8%) ^k	1 (2.1%)	<0.001 ^{b,*}
Liver metastasis	3 (3%)	16 (4.1%)	2 (4.3%)	0.866 ^b
Nonregional LN	6 (6%)	166 (42.5%) ^k	8 (17%)	<0.001 ^{b,*}
Follow-up period (days), median (IQR)	470 (245-730) ^{k,l}	901 (281-2146) ^{k,m}	261 (100-683) ^{l,m}	<0.001 ^{c,*}
Exitus (death)	9 (9%)	184 (47.1%)	34 (72.3%) ^k	<0.001 ^{b,*}
Time from diagnosis to exitus (days), median (IQR)	403 (207.50-7339)	990.50 (220-1826.75)	175 (90.50-405.25) ^k	<0.001 ^{c,*}

ADT: Androgen deprivation therapy; ARPI: Androgen receptor pathway inhibitors; CT: Chemotherapy; CAD: Coronary artery disease; COPD: Chronic obstructive pulmonary disease; CKD: Chronic kidney disease; LN: Lymph nodes; IQR: Interquartile range; a: ANOVA test (post hoc: Bonferroni test); b: Chi-square test; c: Kruskal-Wallis test. k,l,m: Letters indicate statistically significant differences between groups. [†]: This symbol is used to indicate that the percentages represent the distribution of patients according to their local treatment type categories. *, Statistically significant ($p<0.05$)

The distribution of tumor grades also varied significantly by treatment group. Grade 1 disease was most prevalent in the ARPI group (57.5%), whereas grade 3 tumors predominated in the CT group (78.7%) and, to a lesser extent, in the ADT group (36.0%, $p<0.001$). Local treatment was most frequently administered in the ARPI group (28.6%), although this difference did not reach statistical significance ($p=0.063$). Among patients undergoing local treatment, radiotherapy was most commonly delivered in the ADT group (85.0%), whereas surgery was more frequent in the CT group (87.5%, $p<0.001$) (Table 3).

Metastatic disease at diagnosis was significantly more frequent in the ARPI (95.9%) and CT (76.6%) groups than in the ADT group (32.0%; $p<0.001$). Bone metastases were most common in the ARPI group (74.9%), and nonregional lymph node involvement was also significantly higher in this group (42.5%, $p<0.001$). Lung metastases were relatively uncommon in the ADT and CT groups but were observed in 13.8% of ARPI patients ($p<0.001$). No significant differences in the prevalence of liver metastases were observed across groups (Table 3).

The median follow-up duration was significantly longer in the ARPI group (901 days; IQR, 281-2146) than in the ADT (470 days; IQR, 245-730) and CT (261 days; IQR, 100-683) groups ($p<0.001$). Mortality was highest in the CT group (72.3%), followed by the ARPI (47.1%) and ADT (9.0%) groups ($p<0.001$). Furthermore, the median time from diagnosis to death was significantly shorter in the CT group (175 days; IQR, 90.5-405.3) than in the ARPI (990.5 days; IQR, 220-1826.8) and ADT (403 days; IQR, 207.5-7339) groups ($p<0.001$) (Table 3).

When patients were stratified by ARPI treatment type-Abiraterone ($n=133$) versus Enzalutamide ($n=258$)-most clinical and demographic characteristics were comparable between the two groups, with only a few statistically significant differences observed (Table 4). There was no significant difference in mean age at diagnosis (68.51 ± 9.59 vs. 69.87 ± 8.71 years, $p=0.160$). Similarly, family history of prostate cancer was uncommon in both groups (1.5% vs. 3.1%, $p=0.505$). The overall prevalence of comorbidities was similar (59.4% vs. 62.8%, $p=0.513$), with no differences in hypertension (50.4% vs. 53.5%, $p=0.559$), coronary artery disease (26.3% vs. 25.6%, $p=0.875$), chronic obstructive pulmonary disease (1.5% vs. 3.1%, $p=0.505$), or chronic kidney disease (0.8% vs. 0.4%, $p=1.000$). However, diabetes was significantly more prevalent in the Enzalutamide group (26.7% vs. 17.3%, $p=0.037$).

Tumor grade distributions before therapy were not significantly different ($p=0.356$). Grade 1 tumors were slightly more common in the Abiraterone group (62.4% vs. 55%), while grade 3 tumors showed a comparable distribution (14.3% vs. 15.9%).

The vast majority of patients had metastatic disease at baseline, with no difference between groups (96.2% vs. 95.7%, $p=0.812$). Patterns of metastatic spread were also similar: bone metastasis (75.2% vs. 74.8%, $p=0.934$), lung metastasis (13.5% vs. 14.0%, $p=0.909$), liver metastasis (3.0% vs. 4.7%, $p=0.437$), and nonregional lymph node involvement (42.1% vs. 42.6%, $p=0.920$).

Table 4. Comparison of clinical and demographic features by ARPI type

Characteristic	Abiraterone (n=133)	Enzalutamide (n=258)	p
Age at diagnosis (years), mean $\pm$ SD	68.51 $\pm$ 9.59	69.87 $\pm$ 8.71	0.160 ^a
Family history	2 (1.5%)	8 (3.1%)	0.505 ^b
Comorbidities	79 (59.4%)	162 (62.8%)	0.513 ^c
Hypertension	67 (50.4%)	138 (53.5%)	0.559 ^c
CAD	35 (26.3%)	66 (25.6%)	0.875 ^c
Diabetes	23 (17.3%)	69 (26.7%)	0.037 ^{c,*}
COPD	2 (1.5%)	8 (3.1%)	0.505 ^b
CKD	1 (0.8%)	1 (0.4%)	1.000 ^b
Grade before therapy			0.356 ^c
1	83 (62.4%)	142 (55%)	
2	31 (23.3%)	75 (29.1%)	
3	19 (14.3%)	41 (15.9%)	
Metastatic sites	128 (96.2%)	247 (95.7%)	0.812 ^c
Bone metastasis	100 (75.2%)	193 (74.8%)	0.934 ^c
Lung metastasis	18 (13.5%)	36 (14%)	0.909 ^c
Liver metastasis	4 (3%)	12 (4.7%)	0.437 ^c
Nonregional LN	56 (42.1%)	110 (42.6%)	0.920 ^c
Follow-up period (days), median (IQR)	765 (284.50-1780.50)	998.50 (280.50-2296.25)	0.213 ^d
Exitus (death)	52 (39.1%)	132 (51.2%)	0.024 ^{c,*}
Time from diagnosis to exitus (days), median (IQR)	643.50 (179.50-1678)	1016 (237-2128)	0.213 ^d

ARPI: Androgen receptor pathway inhibitors, SD: Standard deviation, CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, CKD: Chronic kidney disease, LN: Lymph nodes, IQR: Interquartile range
a: Independent T test, b: Fisher's exact test, c: Chi-square test, d: Mann-Whitney U test. *: Statistically significant ($p<0.05$)

Median follow-up was longer in the Enzalutamide group (998.5 vs. 765 days), although this did not reach statistical significance ($p=0.213$). Mortality, however, was significantly higher among patients treated with Enzalutamide (51.2% vs. 39.1%, $p=0.024$). Despite this, the median time from diagnosis to death was not statistically different between groups (1016 vs. 643.5 days, $p=0.213$), although numerically longer in the Enzalutamide cohort.

When patients were stratified by medication use ($\geq$ three medications, $n=185$; <3 medications, $n=353$), significant differences were observed in several clinical and demographic characteristics (Table 3). There was no significant difference in mean age at diagnosis between the groups (69.9 ± 10.6 vs. 70.5 ± 11.1 years, $p=0.590$). A positive family history of prostate cancer was significantly less frequent among patients taking ≥ 3 medications (1.6% vs. 2.6%, $p=0.048$).

Tumor grade distribution differed significantly between groups; grade 1 tumors were more prevalent among patients using ≥ 3 medications (65.4% vs. 42.2%), whereas grade 3 tumors were more common in those using <3 medications (34.8% vs. 5.4%, $p<0.001$). The prevalence of comorbidities was markedly higher in the ≥ 3 medications group (80.5% vs. 48.4%, $p<0.001$). Specifically, hypertension (71.9% vs. 33.1%), coronary artery disease (41.1% vs. 11.0%), and diabetes mellitus (35.1% vs. 11.9%) were all significantly more common in the ≥ 3 medications group (all $p<0.001$). No significant differences were observed for COPD or chronic kidney disease.

Metastatic disease was significantly more frequent in the ≥ 3 medications group compared with the < 3 medications group (93.5% vs. 76.5%, $p < 0.001$). Patients in the ≥ 3 medications group also had higher rates of bone metastasis (73.5% vs. 57.2%), lung metastasis (16.8% vs. 7.4%), and nonregional lymph node involvement (44.9% vs. 27.5%) ($p < 0.001$, $p < 0.001$, and $p = 0.001$, respectively).

Regarding treatment type, the majority of patients using ≥ 3 medications received androgen receptor pathway inhibitor (ARPI) therapy (92.4%), while only 7.6% received ADT alone. In contrast, among patients using < 3 medications, 24.4% received ADT alone and 13.3% received CT ($p < 0.001$).

The median follow-up duration was significantly longer in the ≥ 3 medications group compared with the < 3 Medications group (1081 vs. 573 days, $p < 0.001$). Mortality was lower among patients taking ≥ 3 medications (32.4% vs. 47.3%, $p = 0.001$), and the median time from diagnosis to death was significantly prolonged (1254.5 vs. 565 days, $p = 0.001$).

In summary, patients on multiple medications exhibited a higher burden of comorbidities and more advanced metastatic disease but were more likely to have lower tumor grade, experienced longer follow-up, and demonstrated better survival outcomes. These findings suggest that use of three or more medications is common among patients with complex clinical profiles and that treatment and monitoring strategies may need to be tailored to this subgroup (Table 5).

A multivariable Cox proportional hazards regression analysis was conducted to evaluate factors associated with OS (Table 6). The overall model was statistically significant ($p < 0.001$). Older age at diagnosis was independently associated with worse survival, with each additional year increasing the hazard of death by 2.1% (HR=1.021, 95% CI: 1.009-1.034, $p = 0.001$). Tumor grade was significantly associated with survival. Compared to grade 1, patients with grade 2 had a significantly lower hazard of death (HR=0.386, 95% CI: 0.269-0.554, $p < 0.001$), while those with grade 3 had a significantly higher hazard of death (HR=3.757, 95% CI: 2.543-5.551, $p < 0.001$). Medication use of three or more was associated with a nearly two-fold increased hazard of death (HR=1.906, 95% CI: 1.360-2.672, $p < 0.001$). Diabetes mellitus was associated with a 43% reduction in the hazard of death (HR=0.573, 95% CI: 0.405-0.809, $p = 0.002$). Presence of metastasis was associated with a significantly increased hazard of death (HR=2.201, 95% CI: 1.288-3.761, $p = 0.004$). Comorbidity status and bone metastasis were not significantly associated with OS in the adjusted model ($p > 0.05$).

Kaplan-Meier survival analysis demonstrated significant differences in survival times among prostate cancer patients according to treatment type ($p < 0.001$, log-rank test) (Table 7, Figure 1). Patients treated with ADT combined with novel ARPI had the longest mean (2848.8 days) and median (2233 days) survival times. In contrast, patients receiving only CT had the shortest survival times, with a mean of 655.1 days and a median of 277 days. Patients treated with ADT alone had intermediate survival. These results suggest that combining ADT with ARPI may provide a significant survival benefit compared to other treatment modalities.

Table 5. Comparison of clinical and demographic features by medication use

Characteristic	≥ 3 medications (n=185)	< 3 medications (n=353)	p
Age at diagnosis (years), mean $\pm$ SD	69.91 $\pm$ 10.59	70.45 $\pm$ 11.06	0.590 ^a
Family history	3 (1.6%)	18 (2.6%)	0.048 ^{b,*}
Comorbidities	149 (80.5%)	171 (48.4%)	< 0.001 ^{b,*}
Hypertension	133 (71.9%)	117 (33.1%)	< 0.001 ^{b,*}
CAD	76 (41.1%)	39 (11%)	< 0.001 ^{b,*}
Diabetes	65 (35.1%)	42 (11.9%)	< 0.001 ^{b,*}
COPD	5 (2.7%)	12 (3.4%)	0.661 ^b
CKD	4 (2.2%)	3 (0.8%)	0.240 ^c
Grade before therapy			< 0.001 ^{b,*}
1	121 (65.4%) ^k	149 (42.2%)	
2	54 (29.2%)	81 (22.9%)	
3	10 (5.4%)	132 (34.8%) ^l	
Local treatment type	54 (29.2%)	86 (24.4%)	0.226 ^b
Radiotherapy ^y	26 (48.1%)	47 (54.7%)	0.453 ^b
Operation ^y	28 (51.9%)	39 (45.3%)	
Metastatic sites	173 (93.5%)	270 (76.5%)	< 0.001 ^{b,*}
Bone metastasis	136 (73.5%)	202 (57.2%)	< 0.001 ^{b,*}
Lung metastasis	31 (16.8%)	26 (7.4%)	0.001 ^{b,*}
Liver metastasis	6 (3.2%)	15 (4.2%)	0.567 ^b
Nonregional LN	83 (44.9%)	97 (27.5%)	< 0.001 ^{b,*}
Treatment type			< 0.001 ^{b,*}
Only ADT	14 (7.6%)	86 (24.4%) ^k	
ADT with ARPI	171 (92.4%) ^l	220 (62.3%)	
ADT with chemotherapy	0 (0%)	47 (13.3%) ^m	
Follow-up period (days), Median (IQR)	1081 (368.50-2278.50)	573 (218.50-1228.50)	< 0.001 ^{d,*}
Exitus (death)	60 (32.4%)	167 (47.3%)	0.001 ^{b,*}
Time from diagnosis to exitus (days), median (IQR)	1254.50 (576.75-2209.50)	565 (156-1402)	0.001 ^{d,*}

SD: Standard deviation, CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, CKD: Chronic kidney disease, LN: Lymph nodes, ADT: Androgen deprivation therapy, ARPI: Androgen receptor pathway inhibitors, IQR: Interquartile range
a: Independent T test, b: Chi-square test, c: Fisher's exact test, d: Mann-Whitney U test. k,l,m: Letters indicate statistically significant differences between groups. ^y: This symbol is used to indicate that the percentages represent the distribution of patients according to their local treatment type categories. *: Statistically significant ($p < 0.05$)

Table 6. Multivariable cox proportional hazards regression analysis of factors associated with overall survival

Variable	p	HR [Exp (B)]	95% CI for HR
Age at diagnosis (years)	0.001*	1.021	(1.009-1.034)
Grade before therapy (ref: grade 1)			
Grade 2	0.001*	0.386	(0.269-0.554)
Grade 3	0.001*	3.757	(2.543-5.551)
Medication use (≥ 3 vs. < 3 meds)	< 0.001 *	1.906	(1.360-2.672)
Comorbidity (yes vs. no)	0.511	0.901	(0.660-1.230)
Diabetes (yes vs. no)	0.002*	0.573	(0.405-0.809)
Metastasis presence (yes vs. no)	0.004*	2.201	(1.288-3.761)
Bone metastasis (yes vs. no)	0.904	1.021	(0.732-1.424)
Treatment type (ref: ADT alone)			
ADT+chemotherapy	< 0.001 *	6.754	(3.221-14.162)
ADT+ARPI	< 0.001 *	5.703	(2.665-12.201)

HR: Hazard ratio, CI: Confidence interval, ADT: Androgen deprivation therapy, ARPI: Androgen receptor pathway inhibitor, *: Statistically significant ($p < 0.05$)

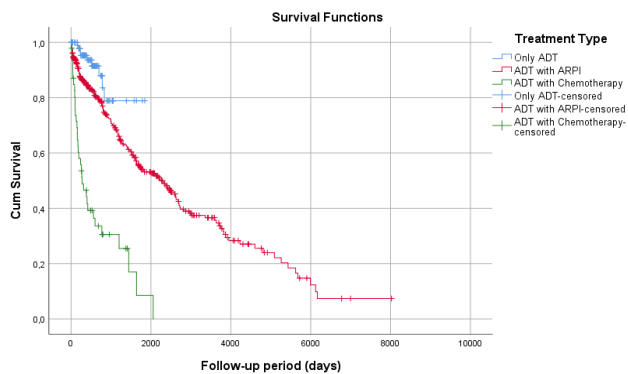


Figure 1. Kaplan-Meier survival curves by treatment type in prostate cancer patients

Kaplan-Meier survival analysis revealed a statistically significant difference in survival times between prostate cancer patients based on medication use ($p < 0.001$, log-rank test) (Table 7, Figure 2). Patients taking three or more medications had a significantly longer mean survival time (3834.8 days; 95% CI: 3167.9-4501.7) and median survival time (2814.0 days; 95% CI: 2099.4-3528.6) compared to those using fewer than three medications, who had a mean survival time of 2195.8 days (95% CI: 1879.9-2511.8) and median survival time of 1510.0 days (95% CI: 1195.1-1824.9). These findings suggest that patients using three or more medications may experience better OS, potentially reflecting differences in patient management or underlying health status.

DISCUSSION

In this multicenter, real-world cohort, we demonstrated that patients with a higher comorbidity burden and concomitant medications (≥ 3 medications) more frequently presented with metastatic disease and had a higher prevalence of bone, lung, and nonregional lymph node metastases. Despite this heavier disease burden, these patients were more likely to receive ARPI therapy rather than chemotherapy or ADT alone and, importantly, achieved longer follow-up, prolonged survival, and lower mortality compared with those taking fewer medications. These findings suggest that ARPIs can be safely and effectively used even in older, more comorbid populations, supporting their feasibility beyond the confines of clinical trial eligibility.¹²

The association between concomitant medications and improved survival outcomes in our cohort likely reflects

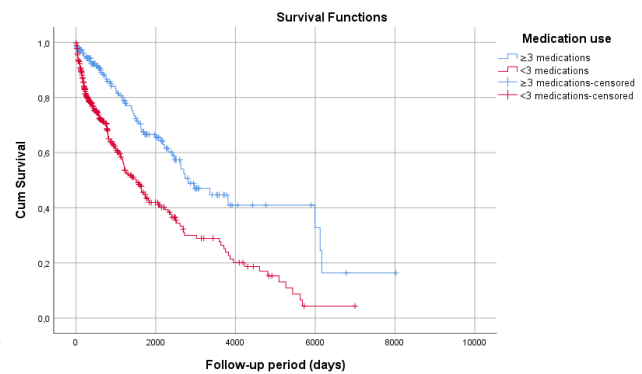


Figure 2. Kaplan-Meier survival curves by medication use in prostate cancer patients

in highly comorbid patients, potentially allowing for greater treatment exposure and greater benefit from ARPI therapy.¹⁵

The preferential use of ARPIs over chemotherapy in patients with concomitant medications is consistent with recent real-world evidence demonstrating that these agents are generally well tolerated, even in frail, elderly, and multimorbid patients.^{16,17} They provide a practical and less toxic alternative to chemotherapy, which may be poorly tolerated in this population due to an increased risk of hematologic and infectious complications.¹⁸

In our real-world cohort comparing Abiraterone (AA) and Enzalutamide (ENZ), baseline clinicodemographic features were broadly similar, with two notable signals: (i) a higher prevalence of diabetes in the ENZ group and (ii) higher crude mortality in ENZ despite comparable disease characteristics and follow-up distribution. Randomized trials establishing the efficacy of each agent in earlier disease settings (ENZAMET for mHSPC and LATITUDE for high-risk mHSPC) demonstrate clear OS benefits versus ADT alone, supporting the oncologic validity of both strategies.^{19,20} However, head-to-head comparative evidence has primarily been derived from observational data. Multiple extensive real-world studies—including a Medicare analysis and nationwide cohorts—tend to show equal or modestly better OS with ENZ versus AA in 1 L mCRPC, particularly among older and comorbid patients.^{21,22}

In contrast, our higher crude mortality in the ENZ arm likely reflects treatment selection patterns rather than intrinsic drug-related differences. In routine oncology practice in Türkiye, frailer or diabetic patients are more frequently prescribed ENZ

Table 7. Kaplan-Meier survival analysis by treatment type and medication use in prostate cancer patients

Group	Subgroup	Mean survival time (days)	95% CI	Median survival time (days)	95% CI
Treatment type	Only ADT	1582.3	1413.2-1751.4	-	-
	ADT with ARPI	2848.8	2497.3-3200.3	2233.0	1718.4-2747.6
	ADT with chemotherapy	655.1	433.1-877.0	277.0	91.0-463.0
	Overall	2742.7	2411.9-3073.5	2110.0	1679.5-2540.5
	Log-rank test: $p < 0.001$ *				
Medication use	≥ 3 medications	3834.8	3167.9-4501.7	2814.0	2099.4-3528.6
	< 3 medications	2195.8	1879.9-2511.8	1510.0	1195.1-1824.9
	Overall	2742.7	2411.9-3073.5	2110.0	1679.5-2540.5
Log-rank test: $p < 0.001$ *					

CI: Confidence interval, ADT: Androgen deprivation therapy, ARPI: Androgen receptor pathway inhibitor, *: Statistically significant ($p < 0.05$)

because Abiraterone requires concomitant corticosteroid use, which may worsen glycemic control or be contraindicated in these individuals. This clinical preference may have enriched the ENZ group for patients with higher baseline comorbidity, leading to apparent mortality differences. Such patterns are consistent with international real-world observations, though additional multicenter data from Turkish practice are needed to clarify these selection dynamics.^{21,22}

The imbalance in diabetes in our data (higher in ENZ) likely reflects treatment selection rather than a drug-induced signal. Abiraterone is given with chronic low-dose corticosteroids, which carry well-known risks of dysglycemia; consistent with this, the EVADE study found a higher risk of new-onset or worsening type 2 diabetes mellitus with AA vs ENZ in mCRPC, despite shorter treatment duration on AA.²³ Case reports and pharmacovigilance also document severe hyperglycemia with AA+prednisone.²⁴ Thus, clinicians may avoid AA in patients with existing diabetes or poor glycemic control, enriching ENZ arms for baseline diabetes—exactly the pattern we observed. This selection dynamic underscores the importance of adjusting for metabolic comorbidities and steroid exposure when comparing outcomes across ARPIs.^{23,24} Finally, cardiovascular-metabolic safety profiles may further influence real-world outcomes. Comparative observational work suggests differential risks of major adverse cardiovascular events across ARPIs, with several datasets indicating greater cardiometabolic liability with steroid-bearing AA versus ENZ. However, findings vary and are sensitive to confounding and analytic choices.^{21,25} Given the broadly similar disease burden across treatment arms, future Turkish real-world analyses should incorporate cause-specific mortality, competing risk models, and detailed comorbidity adjustments to determine whether the observed mortality difference reflects drug effect or patient selection. Prospective registry-based designs with standardized clinical data collection would help address this question more robustly.

Clinically, our findings support considering ARPIs as a first-line systemic therapy option for older patients with high comorbidity burden and multiple drug use, provided that close monitoring and multidisciplinary care are ensured. Individualized treatment plans, medication reconciliation, and proactive management of comorbidities may optimize outcomes and minimize toxicity risks. Importantly, these findings should be interpreted with caution, as they may reflect confounding by indication rather than intrinsic treatment effects. Physicians are to prescribe certain ARPIs to patients who are more closely monitored, have better baseline health status, or possess specific comorbidities that influence treatment choice. Therefore, the observed survival differences may stem from patient selection and closer clinical follow-up rather than the pharmacologic superiority of one agent over another.

Strengths of this study include its multicenter design, relatively large sample size, and comprehensive evaluation of clinical, pathological, and treatment parameters in a real-world setting. Importantly, this analysis addresses a population underrepresented in randomized clinical trials, thereby providing practical insights into everyday oncology practice.

Limitations

This study has several limitations. First, the retrospective, feasibility-focused design may introduce selection bias and preclude causal inference; however, this approach was appropriate for capturing real-world treatment applicability in older, comorbid patients. Second, detailed data on specific drug classes, potential drug-drug interactions, and treatment-related toxicities were not uniformly available, limiting the ability to perform a thorough safety analysis. In addition, data on treatment-related toxicities and quality-of-life outcomes were not consistently collected across centers, particularly among patients receiving multiple concomitant medications. These parameters were therefore not included in the analysis but represent important areas for future research. Third, the severity of comorbidities was not stratified, and specific clinical parameters, such as ECOG performance status, PSA levels, and Gleason score, were inconsistently recorded across centers. However, given that the study's primary aim was to evaluate treatment feasibility rather than prognostic outcomes, these variables were not expected to affect the main findings. Finally, the inclusion of patients from seven centers improves generalizability but may have introduced variability in documentation and follow-up procedures. Additionally, progression-free survival and treatment duration were not analyzed, as the study was not designed to assess treatment efficacy. Future prospective and controlled studies with standardized data collection are warranted to validate and expand upon these findings.

CONCLUSION

In this multicenter, real-world study, we demonstrated that ARPI therapy is feasible and well-tolerated, even in patients with advanced age, a high comorbidity burden, and concomitant medication use. Despite presenting with a higher metastatic burden, these patients achieved longer follow-up, prolonged survival, and lower mortality compared with those on fewer medications or alternative treatments. Our findings support considering ARPIs as a preferred systemic therapy option over chemotherapy or ADT alone in this complex subgroup, provided that close clinical monitoring and multidisciplinary care are implemented. Future prospective studies are warranted to validate these results and to explore quality-of-life and toxicity outcomes in similar high-risk populations.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Şehit Prof. Dr. İlhan Varank Training and Research Hospital Ethics Committee (Date: 19.08.2024, Decision No: 2024/262).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

- Shrestha S, Shrestha S, Khanal S. Polypharmacy in elderly cancer patients: challenges and the way clinical pharmacists can contribute in resource-limited settings. *Aging Med (Milton)*. 2019;2(1):42-49. doi:10.1002/agm2.12051
- Kardas P, Lichwierowicz A, Urbański F, Chudzyńska E, Czech M, Kardas G. Prevalence of chronic polypharmacy in community-dwelling elderly people in Poland: analysis of national real-world database helps to identify high risk group. *Front Pharmacol*. 2021;12:739740. doi:10.3389/fphar.2021.739740.
- Albertsen PC, Moore DF, Shih W, Lin Y, Li H, Lu-Yao GL. Impact of comorbidity on survival among men with localized prostate cancer. *J Clin Oncol*. 2011;29(10):1335-1341. doi:10.1200/JCO.2010.31.2330
- Prathap R, Kirubha S, Rajan AT, Manoharan S, Elumalai K. The increasing prevalence of cancer in the elderly: an investigation of epidemiological trends. *Aging Med (Milton)*. 2024;7(4):516-527. doi:10.1002/agm2.12347
- Feng X, Higa GM, Safarudin F, Sambamoorthi U, Tan X. Potentially inappropriate medication use and associated healthcare utilization and costs among older adults with colorectal, breast, and prostate cancers. *J Geriatr Oncol*. 2019;10(5):698-704. doi:10.1016/j.jgo.2019.01.012
- Nieder C, Mannsäker B, Pawinski A, Haukland E. Polypharmacy in older patients ≥70 years receiving palliative radiotherapy. *Anticancer Res*. 2017;37(2):795-799. doi:10.21873/anticancer.11379
- Saad F, Hamilou Z, Lattouf JB. A drug safety evaluation of enzalutamide to treat advanced prostate cancer. *Expert Opin Drug Saf*. 2021;20(7):741-749. doi:10.1080/14740338.2021.1919620
- Lin YT, Huang YC, Liu CK, Lee TS, Chen M, Chien YN. Treatment-emergent co-morbidities and survival in patients with metastatic castration-resistant prostate cancer receiving abiraterone or enzalutamide. *Front Pharmacol*. 2021;12:669236. doi:10.3389/fphar.2021.669236
- Shaver AL, Nikita N, Sharma S, et al. The safety of abiraterone acetate in patients with metastatic castration-resistant prostate cancer: an individual-participant data meta-analysis based on 14 randomized clinical trials. *Cancers (Basel)*. 2025;17(17):2747. doi:10.3390/cancers17172747
- Shen J, Chowdhury S, Agarwal N, et al. Apalutamide efficacy, safety, and wellbeing in older patients with advanced prostate cancer from phase 3 randomised clinical studies TITAN and SPARTAN. *Br J Cancer*. 2024;130(1):73-81. doi:10.1038/s41416-023-02492-8
- Raval AD, Chen S, Littleton N, Constantinovici N, Goebell PJ. Real-world use of androgen-deprivation therapy intensification for metastatic hormone-sensitive prostate cancer: a systematic review. *BJU Int*. 2025;135(3):408-421. doi:10.1111/bju.16577
- Mir N, Burke O, Yates S, et al. Androgen receptor pathway inhibitors, prostate cancer, and older adults: a global Young International Society of Geriatric Oncology drug review. *Ther Adv Med Oncol*. 2023;15:17588359221149887. doi:10.1177/17588359221149887
- Gaber CE, Okpara E, Abdelaziz AI, et al. Real-world effectiveness and cardiovascular safety of abiraterone versus enzalutamide amongst older patients diagnosed with metastatic castration-resistant prostate cancer. *J Geriatr Oncol*. 2025;16(2):102148. doi:10.1016/j.jgo.2024.102148
- Bolek H, Yazgan SC, Yekedüz E, et al. Androgen receptor pathway inhibitors and drug-drug interactions in prostate cancer. *ESMO Open*. 2024;9(11):103736. doi:10.1016/j.esmoop.2024.103736
- Roviello G, Cappelletti MR, Zanotti L, et al. Targeting the androgenic pathway in elderly patients with castration-resistant prostate cancer: a meta-analysis of randomized trials. *Medicine (Baltimore)*. 2016;95(43):e4636. doi:10.1097/MD.0000000000004636
- Graff JN, Baciarello G, Armstrong AJ, et al. Efficacy and safety of enzalutamide in patients 75 years or older with chemotherapy-naïve metastatic castration-resistant prostate cancer: results from PREVAIL. *Ann Oncol*. 2016;27(2):286-294. doi:10.1093/annonc/mdv542
- George DJ, Ramaswamy K, Yang H, et al. Real-world overall survival with abiraterone acetate versus enzalutamide in chemotherapy-naïve patients with metastatic castration-resistant prostate cancer. *Prostate Cancer Prostatic Dis*. 2024;27(4):756-764. doi:10.1038/s41391-024-00816-0
- Merseburger AS, Dornstauder E, Ohlmann CH, et al. Cardiovascular risks and survival with abiraterone vs enzalutamide in chemotherapy-naïve metastatic castration-resistant prostate cancer in Germany: AVENGER study. *Adv Ther*. 2025;42(4):1919-1934. doi:10.1007/s12325-025-03132-8
- Davis ID, Martin AJ, Stockler MR, et al. Enzalutamide with standard first-line therapy in metastatic prostate cancer. *N Engl J Med*. 2019;381(2):121-131. doi:10.1056/NEJMoa1903835
- Fizazi K, Tran N, Fein L, et al. Abiraterone acetate plus prednisone in patients with newly diagnosed high-risk metastatic castration-sensitive prostate cancer (LATITUDE): final overall survival analysis of a randomised, double-blind, phase 3 trial. *Lancet Oncol*. 2019;20(5):686-700. doi:10.1016/S1470-2045(19)30082-8
- Huang WK, Su PJ, Chen CC, et al. Comparative effectiveness and safety of enzalutamide versus abiraterone in patients with metastatic castration-resistant prostate cancer: a nationwide registry-based cohort study from Taiwan. *J Cancer Res Clin Oncol*. 2025;151(11):284. doi:10.1007/s00432-025-06335-2
- La J, Wang L, Corrigan JK, et al. Abiraterone or enzalutamide for patients with metastatic castration-resistant prostate cancer. *JAMA Netw Open*. 2024;7(8):e2428444. doi:10.1001/jamanetworkopen.2024.28444
- Bahl A, Sodatonou H, Snjider R, et al. Enzalutamide versus abiraterone acetate in the development of new-onset or worsening type 2 diabetes mellitus in patients with metastatic castration-resistant prostate cancer: EVADE study. *World J Urol*. 2024;42(1):584. doi:10.1007/s00345-024-05280-y
- Grace M, Vora T, Estil-Las AA, Arsene C. Severe presentation of de novo type II diabetes in abiraterone and prednisone therapy. *Cureus*. 2025;17(8):e89841. doi:10.7759/cureus.89841
- Lee YHA, Hui JMH, Leung CH, et al. Major adverse cardiovascular events of enzalutamide versus abiraterone in prostate cancer: a retrospective cohort study. *Prostate Cancer Prostatic Dis*. 2024;27(4):776-782. doi:10.1038/s41391-023-00757-0

Evaluation of distribution and revisit rates of pediatric otorhinolaryngologic diseases visits to emergency department and outpatient clinics

 Yeşim Yüksel,  Özer Erdem Gür

Department of Otolaryngology, Antalya Training and Research Hospital, Antalya, Türkiye

Cite this article as: Yüksel Y, Gür ÖE. Evaluation of distribution and revisit rates of pediatric otorhinolaryngologic diseases visits to emergency department and outpatient clinics. *J Health Sci Med.* 2025;8(6):1164-1170.

Received: 04.09.2025

Accepted: 17.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aims to evaluate the distribution of acute pediatric otorhinolaryngologic (ORL) diseases and their revisit rates across three clinical departments-pediatric emergency departments, pediatric outpatient clinics, and otorhinolaryngology outpatient clinics-within a tertiary care hospital. The objective is to identify triage patterns that can optimize healthcare delivery and inform appropriate clinic referrals for common pediatric ORL conditions.

Methods: A retrospective cross-sectional study was conducted between January 2014 and December 2024. Patients under the age of 18 who presented to the specified clinics with acute ORL diseases were included based on ICD-10 diagnostic codes. Data from the years 2020-2022 were excluded due to the COVID-19 pandemic. Patient counts and revisit frequencies were obtained from hospital records with support from the Information Technology department and approval from the institutional ethics committee.

Results: The distribution of acute ORL diagnoses showed significant variation across clinics. Pharyngitis, tonsillitis, multiple and unspecified upper respiratory tract infections (URTI), and allergic rhinitis were the most common diagnoses, while nasopharyngitis, laryngitis, croup, and otitis externa were less frequently observed. Pediatric emergency department had higher rates of pharyngitis and unspecified URTI, whereas pediatric outpatient clinics saw more allergic rhinitis and sinusitis cases. The otorhinolaryngology clinic had elevated rates of sinusitis, otitis media (both suppurative and non-suppurative), and allergic rhinitis. Revisit rates also differed: emergency departments had the highest for pharyngitis and URTI; pediatric clinics led in allergic rhinitis revisits; and otorhinolaryngology clinic had the highest revisit rates for sinusitis and otitis media.

Conclusion: Patterns of presentation and revisit for acute pediatric ORL diseases differ significantly across clinical departments. Emergency departments are frequently utilized for conditions that could be more appropriately managed in outpatient settings, potentially delaying access to specialist care. Accurate diagnosis and timely referral are important for preventing complications, especially in ORL diseases such as otitis media and sinusitis. These findings underscore the need for improved triage and education to optimize resource utilization and patient outcomes in pediatric ORL care.

Keywords: Otorhinolaryngologic disease, pediatric, rhinosinusitis, otitis media

INTRODUCTION

Pediatric otorhinolaryngologic (ORL) diseases are among the most frequent reasons for emergency department and outpatient clinic visits. Conditions such as acute otitis media, tonsillitis, and sinusitis are among the commonest childhood disorders and important reasons for parents and children to visit the physicians/otolaryngologists.¹⁻³

The symptoms presentation and initial management of pediatric ORL disorders in children is a significant factor in the outcome of these disorders. It is a well-established fact that initial consultations typically occur with general practitioners, emergency specialist or pediatricians rather than ORL specialists. Many ORL diseases, despite having distinct clinical diagnoses, share overlapping symptoms

such as high fever, difficulty swallowing, nasal discharge, ear pain, headache, and sore throat. These similarities can lead to misdiagnosis, especially in emergency services and primary care settings, resulting in delayed referrals to ORL specialists. In children, acute ORL diseases such as otitis media, tonsillitis, and sinusitis may spread to intracranial structures and the orbital region due to delays in diagnosis and treatment, potentially leading to complications associated with high morbidity and mortality.⁴⁻⁶

Emergency departments are units that provide uninterrupted emergency healthcare services and perform urgent medical interventions. However, this primary function of emergency departments is disrupted due to the large number of patients

Corresponding Author: Yeşim Yüksel, yesimgedikli@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

who present with non-urgent conditions. The main reasons for this are failure to seek treatment in accordance with the definition of an emergency, and the use of emergency departments as primary healthcare units or specialist outpatient clinics.⁷ In addition to actual ORL emergencies such as retropharyngeal abscess, epiglottitis, mastoiditis, peritonsillar abscess, epistaxis, and foreign bodies in the ear or nose, all diseases affecting the ear, nose and throat region are treated in the emergency department.⁶ Particularly in diseases such as acute otitis media and acute rhinosinusitis, the correct diagnosis by specialists in the relevant field using the necessary devices and equipment not only ensures that the necessary treatment is administered but also contributes to preventing negative outcomes such as the disease becoming chronic and the development of complications.^{5,8,9}

Understanding the distribution and revisit patterns of pediatric ORL cases across different clinical departments is essential for improving early diagnosis and optimizing healthcare delivery. This study aims to determine the distribution of the most common acute pediatric ENT diseases in the pediatric emergency department, pediatric diseases, and otorhinolaryngology outpatient clinics of a tertiary hospital, and to compare the revisits rates in the emergency department and outpatient clinics.

METHODS

This retrospective cross-sectional study evaluates patients who presented to the pediatric emergency department, pediatric outpatient clinics, and otorhinolaryngology outpatient clinics between January 2014 and December 2024 due to acute ORL diseases. The study was approved by the Antalya Training and Research Hospital Scientific Researches Ethics Committee (Date: 03.07.2025, Decision No: 11/23). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

All patients included in the study were selected using the diagnosis codes from the ICD-10 version of the International Statistical Classification of Diseases and Related Health Problems (ICD) system, published by the World Health Organisation. Data from 2020, 2021, and 2022, the years in which the coronavirus disease (COVID-19) pandemic was reported by the WHO as active, and patients aged 18 years and older were excluded from the study. Additionally, acute ORL disease diagnosis codes recorded in clinic visits outside the clinics included in the study, visits in which the diagnosis was not made by the relevant specialist physicians, patients who received more than one acute ORL disease diagnosis code during their visit, and patients whose visit included an additional diagnosis code other than the acute ORL disease diagnosis code were excluded from the study. Visits made to the relevant clinics within 10 days after a previous visit were not included in the study.

All patients under the age of 18 who were diagnosed by specialist physicians with the diseases corresponding to the ICD-10 diagnosis codes listed in [Table 1](#) were included in the study. In ICD-10 diagnosis codes, the digits following the decimal point represent more specific subcategories or clinical details of a disease or condition. To minimize the potential impact

of inaccurate differential diagnoses and to ensure optimal evaluation, these subcategories have been considered within the framework of the general category code of the disease, as presented in [Table 1](#). The number of patients diagnosed with the selected ICD-10 codes by relevant specialists, along with the number of repeat visits made by the same patients to the same department, were obtained from hospital records by the Information Technology departments after receiving the necessary permissions. The number of patients and repeat visits corresponding to each diagnosis code were recorded separately for the pediatric emergency department, pediatric outpatient clinic, and otorhinolaryngology outpatient clinic, and were subsequently analyzed through comparative evaluation.

Table 1. ICD-10 codes and disease descriptions included in the study

J00	Nasopharyngitis
J01	Sinusitis
J02	Pharyngitis
J03	Tonsillitis
J04	Laryngitis and tracheitis
J05	Croup and epiglottitis
J06	Multiple and Unspecified URTI
J30	Vasomotor and allergic rhinitis
H60	Otitis externa
H65	Non-suppurative OM
H66	Suppurative and unspecified OM
URTI: Upper respiratory tract infection, OM: Otitis media	

Statistical Analysis

The statistical analysis of the obtained data was performed using the SPSS for Windows 11.5 software package. The Chi-square test and one-sample Chi-square test were used for statistical analysis. Descriptive values were expressed as frequencies and percentages. The statistical significance threshold was set at 0.05.

RESULTS

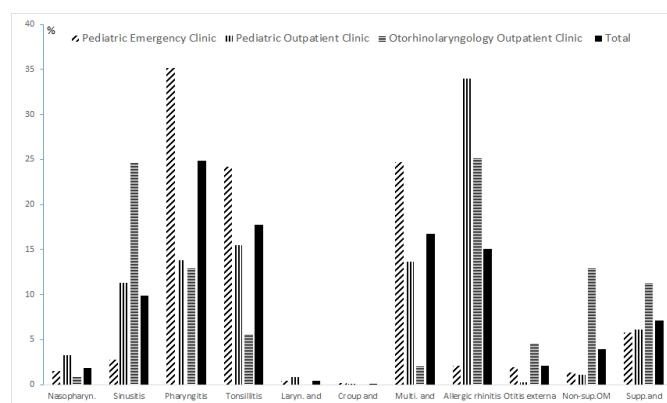
When examining the distribution of total patient numbers and rates attending the emergency department, pediatric clinics and otorhinolaryngology outpatient clinics, it was found that the distribution of acute ORL disease diagnosis groups was not similar. Nasopharyngitis, laryngitis and tracheitis, croup, and epiglottitis, and otitis externa were observed at low rates, while pharyngitis, tonsillitis, multiple and unspecified acute upper respiratory tract infections (URTI), and allergic rhinitis were found to be high. This difference in distribution was found to be statistically significant ([Table 2, Figure 1](#)).

Patient numbers were provided separately for each clinic and diagnosis, and the proportions of these numbers among all patients in the same clinic were evaluated ([Table 3](#)). For each separate clinic, differences were observed in the number of patients specific to the diagnosis, their proportions, and the total number of visits. In the pediatric emergency department, it was observed that the distribution of patients with different acute ORL disease diagnoses was not similar. Pharyngitis,

Table 2. The number of patients per diagnosis, the proportional distribution among all patients, and the number of diagnosis-specific visits

Acute ORL disease (2014-2019, 2023-2024)	Total patient		Total outpatient visit#
	n	%	
Nasopharyngitis	12920	1.83	13391
Sinusitis	69798	9.87	78369
Pharyngitis	175664	24.85	210099
Tonsillitis	125447	17.75	140778
Laryngitis and tracheitis	3123	0.44	3207
Croup and epiglottitis	737	0.11	763
Multiple and unspecified URTI	118630	16.78	137262
Allergic rhinitis	106822	15.11	134293
Otitis externa	15051	2.13	15685
Non-suppurative OM	28083	3.97	30683
Suppurative and unspecified OM	50614	7.16	55816
Total	706889	100	820346
p	0.001		

ORL: Otorhinolaryngologic, URTI: Upper respiratory tract infection, OM: Otitis media

**Figure 1.** Diagnosis-specific patient ratios for each clinic and for the total patient cohort

tonsillitis, multiple and unspecified URTI, and suppurative and unspecified OM were found to be higher in proportion than other acute ORL disease patients presenting to the pediatric emergency department. For patients attending the pediatric outpatient clinic, sinusitis, pharyngitis, tonsillitis, multiple and unspecified URTI, and allergic rhinitis were found to be more prevalent, while the distribution of other diagnosis groups was similar. For the otorhinolaryngology outpatient clinic, it was determined that the patient rates for sinusitis, pharyngitis, tonsillitis, allergic rhinitis, non-suppurative OM, and suppurative and unspecified OM were high, while the rates for other diagnostic groups were low and similar to each other (**Table 3, Figure 1**).

It has been determined that the number and rate of visits to the emergency department, pediatric clinic, and otorhinolaryngology outpatient clinic differ between disease groups. The number and rate of outpatient visits for each clinic, by separate diagnosis groups, are given in **Table 4**. The most common diagnoses in the pediatric emergency department were pharyngitis, multiple and unspecified URTI, and tonsillitis. In the pediatric outpatient clinic, the most common diagnoses were allergic rhinitis, tonsillitis, pharyngitis, multiple and unspecified URTI, and sinusitis. The most common diseases presenting at the otorhinolaryngology outpatient clinic were identified as Allergic rhinitis, Sinusitis, Non-suppurative OM, Pharyngitis, and Suppurative and unspecified OM.

Table 5 presents the revisit rates for each acute ORL disease diagnosis group. Inter-clinic comparison revealed no significant differences for the diagnoses of nasopharyngitis, laryngitis and tracheitis, and croup and epiglottitis. In contrast, significant differences were observed between clinics for the diagnoses of sinusitis, pharyngitis, multiple and unspecified URTI, allergic rhinitis, otitis externa, and suppurative and unspecified OM. The emergency department was identified as the clinic with the highest rate

Table 3. The number of patients specific to the clinic and diagnosis they visit, along with the proportional distribution of these patients among all patients of the clinic

Acute ORL disease (2014-2019, 2023-2024)	Pediatric emergency clinic total patient		Pediatric outpatient clinic total patient		Otorhinolaryngology outpatient clinic total patient	
	n	%	n	%	n	%
Nasopharyngitis	5751	1.54	5637	3.31	1532	0.94
Sinusitis	10226	2.74	19333	11.34	40239	24.58
Pharyngitis	130990	35.15	23536	13.81	21138	12.91
Tonsillitis	90018	24.16	26407	15.49	9022	5.51
Laryngitis and tracheitis	1501	0.40	1464	0.86	158	0.10
Croup and epiglottitis	657	0.18	80	0.05	0	0
Multiple and unspecified URTI	92053	24.70	23245	13.64	3332	2.03
Allergic rhinitis	7794	2.09	57946	33.99	41082	25.09
Otitis externa	7122	1.91	485	0.28	7444	4.55
Non-suppurative OM	4955	1.33	1869	1.10	21259	12.98
Suppurative and unspecified OM	21599	5.80	10482	6.15	18533	11.32
Total	372666	100	170484	100	163739	100
p	0.001		0.001		0.001	

ORL: Otorhinolaryngologic, URTI: Upper respiratory tract infection, OM: Otitis media

Table 4. Clinic-level distribution and comparative analysis of total patient visits and visit rates by specific diagnoses

Acute ORL disease (2014-2019, 2023-2024)	Pediatric emergency department visit (1)		Pediatric outpatient clinic visit (2)		Otorhinolaryngology outpatient clinic visit (3)	
	#	%	#	%	#	%
Nasopharyngitis	5882	1.36	5952	2.99	1557	0.82
Sinusitis	10504	2.43	20735	10.40	47130	24.95
Pharyngitis	161010	37.26	26018	13.06	23071	12.21
Tonsillitis	101245	23.43	29574	14.84	9959	5.27
Laryngitis and tracheitis	1539	0.36	1507	0.76	161	0.09
Croup and epiglottitis	682	0.16	81	0.04	0	0.00
Multiple and unspecified URTI	108106	25.02	25746	12.92	3410	1.80
Allergic rhinitis	8024	1.86	75807	38.04	50462	26.71
Otitis externa	7224	1.67	493	0.25	7968	4.22
Non-suppurative OM	5250	1.21	1999	1.00	23434	12.40
Suppurative and unspecified OM	22673	5.25	11373	5.71	21770	11.52
Total	432139	100.0	199285	100.0	188922	100.00
p	0.001					

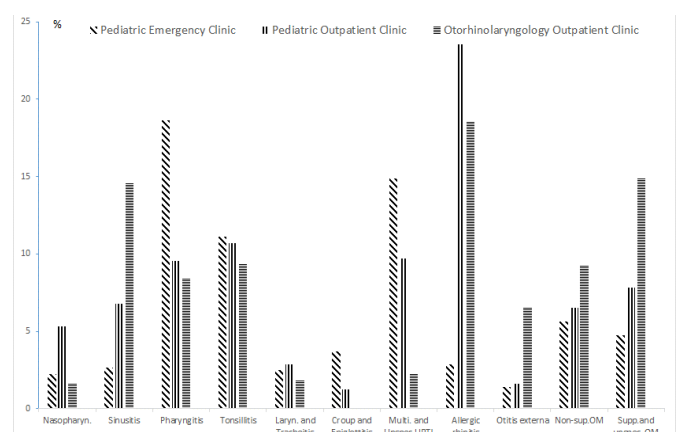
ORL: Otorhinolaryngologic, URTI: Upper respiratory tract infection, OM: Otitis media

Table 5. Diagnosis-specific patient revisit rates by clinic of initial visit and corresponding comparative analysis

Acute ORL disease (2014-2019, 2023-2024)	Pediatric emergency department (1)	Pediatric outpatient clinic (2)	Otorhinolaryngology outpatient clinic (3)	p	Statistically significant groups
	%	%	%		
Nasopharyngitis	2.23	5.29	1.61	0.126	---
Sinusitis	2.65	6.76	14.62	0.001	1-2-3
Pharyngitis	18.64	9.54	8.38	0.001	1-2-3
Tonsillitis	11.09	10.71	9.41	0.001	1-3, 2-3
Laryngitis and tracheitis	2.47	2.85	1.86	0.664	----
Croup and epiglottitis	3.67	1.23	0.0	0.254	----
Multiple and unspecified URTI	14.85	9.71	2.29	0.001	1-2-3
Allergic rhinitis	2.87	23.56	18.59	0.001	1-2-3
Otitis externa	1.41	1.62	6.58	0.001	1-2-3
Non-suppurative OM	5.62	6.50	9.28	0.001	1-3, 2-3
Suppurative and unspecified OM	4.74	7.83	14.87	0.001	1-2-3

ORL: Otorhinolaryngologic, URTI: Upper respiratory tract infection, OM: Otitis media

of repeat visits, particularly for Pharyngitis and Multiple and Unspecified URTI diagnoses, followed by the pediatrics and otorhinolaryngology outpatient clinics. The highest rate of repeat visits for Allergic rhinitis diagnoses was observed in the pediatrics outpatient clinic, which was higher than that of the otorhinolaryngology outpatient clinic. The frequency of repeat visits for sinusitis, otitis externa, and suppurative and unspecified OM patients was found to be, in order, the otorhinolaryngology outpatient clinic, the pediatrics outpatient clinic, and the pediatric emergency department. Furthermore, the rate of repeat visits for tonsillitis diagnosis was found to be lower in the otorhinolaryngology outpatient clinic compared to the pediatric emergency department and pediatric outpatient clinic; however, it was found to be higher for non-suppurative OM diagnosis (Table 5, Figure 2).

**Figure 2.** Diagnosis-specific patient revisit rates based on the clinic of initial visit

DISCUSSION

Evaluating the clinical preferences and number of visits related to acute pediatric ORL diseases is valuable for identifying healthcare utilization patterns and planning service delivery accordingly. In this study, the profile of patients presenting to the pediatric emergency department aligns with complaint-based visits due to acute symptoms and is closely associated with parental concerns regarding the sudden onset and severity of symptoms. Furthermore, the assessment of initial and repeat visits to pediatric and otorhinolaryngology outpatient clinics reflects families' conscious choices in the treatment process and the feasibility of implementing targeted clinical interventions. Although the ICD-10 diagnostic codes have been reported to have certain limitations such as coding accuracy, diagnostic specificity and consistency, their use is nevertheless considered essential because the large sample size enables standardization in epidemiological studies and health system planning, thereby allowing for systematic analysis.¹⁰⁻¹² Conducted in a tertiary care center using ICD-10 diagnostic codes and a large case series, this study provides an opportunity to analyze healthcare utilization habits and evaluate which diseases prompt visits to which clinics and how frequently.

Acute ORL diseases were seen at different frequencies in the clinics visited. The high rates of pharyngitis, tonsillitis, multiple and unspecified URTI, and allergic rhinitis diagnoses in all clinics indicate that these diseases are common in childhood.^{2,3,13-15} In contrast, the absence of significant differences between clinics in the diagnoses of laryngitis and tracheitis, croup and epiglottitis can be explained by the relatively rarer occurrence of these diseases and their similar management.^{13,15}

The high rates of pharyngitis, tonsillitis, and multiple and unspecified URTI in the pediatric emergency department in particular suggest that more acute and symptomatic cases are evaluated in this clinic. This is consistent with the literature, which indicates that emergency departments are generally preferred due to sudden onset and severe symptoms.^{1-3,13,14}

The number and proportion of patients diagnosed with allergic rhinitis are higher in the pediatric outpatient clinic than in the otorhinolaryngology outpatient clinic. Conversely, the number and proportion of patients diagnosed with sinusitis and otitis media are higher in the otorhinolaryngology outpatient clinic compared to the pediatrics outpatient clinic. As determined in our study, the high incidence of allergic rhinitis in pediatric clinics is due to the existence of specific specialist clinics, such as pediatric allergy and immunology clinics, in tertiary hospitals for these patients.^{16,17} In the otorhinolaryngology outpatient clinic, in addition to allergic rhinitis, the high number of patients with sinusitis and non-suppurative and suppurative OM indicates that more specific and follow-up-requiring cases are handled in this clinic. The results of our study regarding the number and proportion of patients attending otorhinolaryngology clinics are consistent with the literature.^{4,9,13,15,18-20} This supports the view that specialist clinics offer a more detailed assessment in terms of diagnosis and treatment planning.

Revisit rates exhibit significant differences based on diagnosis and clinical setting, indicating that disease management processes and patient follow-up vary across clinics. There are two primary reasons for revisits: one is complaint-based visits due to acute symptoms, and the other is scheduled visits. Notably, the highest revisit rates have been observed in pediatric emergency departments for diagnoses such as pharyngitis and multiple/unspecified upper respiratory tract infections (URIs). This is primarily attributed to complaint-based visits due to acute symptoms.²¹ In pediatric outpatient clinics, diseases with high revisit rates, excluding allergic rhinitis, show similarities to those observed in pediatric emergency departments. In contrast, diseases with high revisit rates in otorhinolaryngology outpatient clinics differ substantially from those in pediatric emergency departments and outpatient clinics. This discrepancy may be due to the inclusion of scheduled visits in otorhinolaryngology outpatient clinics, in addition to complaint-based visits due to acute symptoms.

The differential diagnosis of acute ORL diseases may not be adequately performed in emergency department settings. Diagnosing pediatric rhinosinusitis is challenging because the symptoms often resemble those of other URTI, allergic rhinitis, and adenoiditis.²² In addition to the difficulty in expressing symptoms in the pediatric age group, difficulties in diagnosis are caused by factors such as children's inability to tolerate nasal endoscopy, which supports the diagnosis of rhinosinusitis, and the limited use of computed tomography due to radiation exposure.²³ Particularly in diseases such as sinusitis, tonsillitis, suppurative and non-suppurative otitis media, detailed examination in otorhinolaryngology outpatient clinics and endoscopic evaluation in appropriate patients are much more valuable for accurate diagnosis and treatment planning than examination in emergency department conditions and pediatric outpatient clinics where endoscopic examination is not possible are much more valuable in terms of accurate diagnosis and treatment planning compared to examinations performed in emergency department settings and pediatric clinics where endoscopic examination is not possible. This enables the optimal follow-up and treatment of acute ORL diseases that carry a risk of chronicity and/or may lead to serious complications, such as sinusitis and suppurative otitis media.^{5,8,13,22-24} In our study, the high rate of repeat visits to the otorhinolaryngology outpatient clinic by patients with sinusitis, suppurative and non-suppurative otitis media also supports this situation. In the otorhinolaryngology outpatient clinic, it is important that the follow-up process is more regular and long-term in order to be able to decide on surgical treatment for these diseases when necessary.⁸ In contrast, the low rate of repeat visits for tonsillitis, particularly in the otorhinolaryngology outpatient clinic, may indicate the effectiveness of the treatment and follow-up protocols applied in this clinic. Among the acute ORL diseases included in our study, the highest rates of repeat visits were observed in allergic rhinitis. Particularly in allergic rhinitis, difficulties in managing symptoms and the need for long-term follow-up lead to repeated patient visits to pediatric and otorhinolaryngology outpatient clinics. The frequency of repeat visits by allergic rhinitis patients is expected due to the

condition affecting all pediatric age groups and its associated diseases.^{17,25}

Limitations

This study is a healthcare utilization analysis focused on disease burden, clinical preferences, and revisit behaviors, based on the internationally recognized ICD-10 diagnosis codes. Although not a prevalence or incidence study, this retrospective cross-sectional study reveals the clinical visit behaviors and patient orientation patterns of pediatric acute ORL disease cases at a tertiary care hospital, supported by a large case series spanning several years. However, since the study is based on selected ICD-10 diagnostic codes, the distribution of patients and revisit rates is inherently dependent on the accuracy of clinical assessments and diagnostic precision by of the respective clinicians. In clinical practice, these diagnoses are used interchangeably or may overlap. For example, the differential diagnosis between nasopharyngitis and pharyngitis may not always be clearly established, and due to symptomatic similarities, clinicians may have entered the ICD-10 code for “Multiple and Unspecified URTI” instead of more specific diagnoses such as nasopharyngitis, pharyngitis, or sinusitis. The potential inconsistency in ICD-10 code selection and limitations in differential diagnosis could have occurred similarly or differently across all physicians in the pediatric emergency department, pediatric outpatient clinics, and otorhinolaryngology outpatient clinics. However, due to the retrospective nature of the study, it is not possible to determine this. Repeat visits to the pediatric emergency department for the same ICD-10 diagnosis code generally represent complaint-based visits due to acute symptoms, whereas visits to pediatric and otorhinolaryngology outpatient clinics may be either complaint-based or scheduled visits. Additionally, the fact that the data were obtained from a single center, and that demographic and socio-economic variables, follow-up intervals, and treatment plans were not included in the analysis, are also among the limitations of this study.

CONCLUSION

The findings obtained in our study are epidemiologically significant due to the large sample size. This study has revealed marked differences in the distribution of acute ORL diseases across clinics and in the rates of repeat visits. These differences are shaped by the functional roles of the clinics, patient profiles, and the nature of the diseases. It is recommended that clinic-specific follow-up protocols be developed, family education and symptom management strategies be strengthened to reduce revisit rates. Taking these data into account in the planning of healthcare services is important for the effective use of resources and increasing patient satisfaction.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Antalya Training and Research Hospital Scientific Researches Ethics Committee (Date: 03.07.2025, Decision No: 11/23).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Yoong SYC, Ang PH, Chong SL, et al. Common diagnoses among pediatric attendances at emergency departments. *BMC Pediatr*. 2021;21(1):172:1-6 doi:10.1186/s12887-021-02646-8
2. Signorelli LG, Mendes Ede A. Prevalence of otorhinolaryngologic diagnoses in the pediatric emergency room. *Int Arch Otorhinolaryngol*. 2013;17(1):10-3. doi:10.7162/S1809-9772013000100002
3. Merino-Galvez E, Gomez-Hervas J, Perez-Mestre D, Llamas-Peiro JM, Perez-Gil E, Belda-Palazon M. Epidemiology of otorhinolaryngologic emergencies in a secondary hospital: analysis of 64,054 cases. *Eur Arch Otorhinolaryngol*. 2019;276(3):911-917. doi:10.1007/s00405-019-05331-w
4. Gupta V, Gupta A. Pattern of paediatric ear, nose and throat disorders in a district hospital. *Int J Otorhinolaryngol Head Neck Surg*. 2019;5(2):403-407.
5. Torretta S, Drago L, Marchisio P, et al. Review of systemic antibiotic treatments in children with rhinosinusitis. *J Clin Med*. 2019;8(8):1162. doi:10.3390/jcm8081162
6. Stoner MJ, Dulaaurier M. Pediatric ENT emergencies. *Emerg Med Clin North Am*. 2013;31(3):795-808. doi:10.1016/j.emc.2013.04.005
7. Yıldız Y, Kanburoglu M. Health service delivery quality and patient satisfaction in pediatric emergency department. *J Pediatr Emerg Intensive Care Med*. 2021;8:7-14 doi:10.4274/cayd.galenos.2020.08379
8. Jamal A, Alsabea A, Tarakme M, Safar A. Etiology, diagnosis, complications, and management of acute otitis media in children. *Cureus*. 2022;14(8):e28019. doi:10.7759/cureus.28019
9. Çolak M, Güver Çalapkulu İD, Taşdemir A, Yağmur AR, Akbal Çufalı Ş, Gülhan Yaşar N. Why pediatricians need to consult an otolaryngologist: analysis of 3774 patients. *Osmangazi J Med*. 2022;44(6):799-806.
10. Nelson SJ, Yin Y, Trujillo Rivera EA, et al. Are ICD codes reliable for observational studies? Assessing coding consistency for data quality. *Digit Health*. 2024;10:20552076241297056. doi:10.1177/20552076241297056
11. Zarei J, Golpira R, Hashemi N, et al. Comparison of the accuracy of inpatient morbidity coding with ICD-11 and ICD-10. *Health Inf Manag*. 2025;54(1):14-24. doi:10.1177/18333583231185355
12. Peng M, Eastwood C, Boxill A, et al. Coding reliability and agreement of International Classification of Disease, 10th revision (ICD-10) codes in emergency department data. *Int J Popul Data Sci*. 2018;3(1):445. doi:10.23889/ijpds.v3i1.445
13. Ciprandi G, Passali FM, Passali D. The impact of infections in the otolaryngologist's practice. *Acta Biomed*. 2022;93(5):e2022215. doi:10.23750/abm.v93i5.12630
14. Weintraub B. Upper respiratory tract infections. *Pediatr Rev*. 2015;36(12):554-556. doi:10.1542/pir.36-12-554
15. Dzieciołowska-Baran E, Gawlikowska-Sroka A, Mularczyk M. Diseases of the upper respiratory tract in preschool and school age children in ambulatory ear nose throat practice. *Adv Exp Med Biol*. 2015;873:35-41. doi:10.1007/5584_2015_132

16. Şimşek Y, Yılmaz Ö, Yüksel H. Allergic rhinitis. *Asthma Allergy Immunology* 2018;16:059-069. doi:10.21911/aa.373
17. Chrysouli K, Theodorakopoulos C, Saratsiotis A, et al. Allergic rhinitis in children: an underestimated disease. *Indian J Otolaryngol Head Neck Surg.* 2024;76(2):1759-1764. doi:10.1007/s12070-023-04402-z
18. Castelli Gattinara G, Bergamini M, Simeone G, et al. Antibiotic treatment of acute and recurrent otitis media in children: an Italian intersociety Consensus. *Ital J Pediatr.* 2025;51(1):50. doi:10.1186/s13052-025-01894-z
19. Tall H, Bah FY, Nasser T, Sambou A, Diallo BK. Ear, nose and thorax disorders in pediatric patients at a rural hospital in Senegal. *Int J Pediatr Otorhinolaryngol.* 2017;96:1-3. doi:10.1016/j.ijporl.2017.02.019
20. Uijen JH, Bindels PJ, Schellevis FG, van der Wouden JC. ENT problems in Dutch children: trends in incidence rates, antibiotic prescribing and referrals 2002-2008. *Scand J Prim Health Care.* 2011;29(2):75-79. doi:10.3109/02813432.2011.569140
21. Corsello A, Milani GP, Picca M, et al. Recurrent upper respiratory tract infections in early childhood: a newly defined clinical condition. *Ital J Pediatr.* 2024;50(1):30. doi:10.1186/s13052-024-01600-5
22. Siddiqui Z, Tahiri M, Gupta A, Kin Nam RH, Rachmanidou A. The management of paediatric rhinosinusitis. *Int J Pediatr Otorhinolaryngol.* 2021;147:110786. doi:10.1016/j.ijporl.2021.110786
23. Gilani S, Shin JJ. The burden and visit prevalence of pediatric chronic rhinosinusitis. *Otolaryngol Head Neck Surg.* 2017;157(6):1048-1052. doi:10.1177/0194599817721177
24. Hullegie S, Venekamp RP, van Dongen TMA, et al. Prevalence and antimicrobial resistance of bacteria in children with acute otitis media and ear discharge: a systematic review. *Pediatr Infect Dis J.* 2021;40(8):756-762. doi:10.1097/INF.0000000000003134
25. Vardouniotis A, Doulapsi M, Aoi N, Karatzanis A, Kawauchi H, Prokopakis E. Local allergic rhinitis revisited. *Curr Allergy Asthma Rep.* 2020;20(7):22. doi:10.1007/s11882-020-00925-5

Comparison of operative and nonoperative treatment of midshaft clavicle fractures: early superiority of surgery with comparable long-term outcomes

 Zafer Güneş,  Eralp Erdoğan

Department of Orthopedics and Traumatology, Ankara Training and Research Hospital, University of Health Sciences, Ankara, Türkiye

Cite this article as: Güneş Z, Erdoğan E. Comparison of operative and nonoperative treatment of midshaft clavicle fractures: early superiority of surgery with comparable long-term outcomes. *J Health Sci Med.* 2025;8(6):1171-1175.

Received: 14.08.2025

Accepted: 19.10.2025

Published: 25.10.2025

ABSTRACT

Aims: Midshaft clavicle fractures are common injuries in young, active adults. Although nonoperative management has traditionally been preferred, recent evidence suggests that surgical fixation may yield faster recovery and lower rates of nonunion. This study aimed to compare the clinical and radiological outcomes of operative versus nonoperative treatment of midshaft clavicle fractures.

Methods: A retrospective cohort of 98 patients (34 surgically, 64 conservatively treated) with Robinson type 2B1-2B2 midshaft clavicle fractures was analyzed. Patients aged 18-65 years and followed for at least 12 months were included. Surgical indications comprised displacement >2 cm, comminution, skin tenting, or neurovascular compromise. Radiographic union, Constant-Murley, and Visual Analog Scale (VAS; 0-100 mm) scores were compared at 3, 6, and 12 months. Kaplan-Meier survival analysis was used to evaluate time to union, with intergroup comparison by log-rank test.

Results: Mean union time was 16.4±3.2 weeks in the surgical group and 24.3±4.6 weeks in the conservative group ($p<0.001$). Kaplan-Meier analysis showed significantly faster union in the surgical cohort (log-rank $\chi^2=17.46$, $p<0.001$). Constant-Murley scores were significantly higher and VAS scores lower at 3 and 6 months after surgery ($p<0.001$ for both), but differences were not significant at 12 months. Complications occurred in 14.7% of surgical and 17.2% of conservative cases ($p=0.260^*$). Nonunion (14%) and malunion (3%) were observed only in the conservative group.

Conclusion: Operative fixation of displaced midshaft clavicle fractures provides faster radiologic healing and better early functional outcomes, while long-term results are comparable to conservative management. Surgical treatment should be preferred in young, active patients or markedly displaced fractures, whereas nonoperative care remains suitable for minimally displaced, low-demand cases.

Keywords: Clavicle fracture, midshaft, open reduction and internal fixation, conservative treatment, Kaplan-Meier, functional outcome, union time

INTRODUCTION

Accounting for 2%-5% of all fractures,¹ clavicle fractures are common injuries, particularly among young and active adults. The midshaft region is the most frequently affected segment due to its slender shape and limited muscular coverage.^{2,3} Displacement is common, resulting from the opposing pull of the sternocleidomastoid, pectoralis major, and deltoid muscles.⁴

Despite substantial advances in orthopedic trauma care, the optimal treatment strategy for midshaft clavicle fractures remains controversial.^{5,6} Historically, nonoperative management was preferred because of the bone's high remodeling potential and satisfactory long-term outcomes. Conservative treatment aims to relieve pain and maintain acceptable alignment using a figure-of-eight brace or arm sling. However, increasing evidence indicates that non-

surgical management can lead to higher rates of nonunion, malunion, and residual shoulder dysfunction compared to surgery.⁷⁻⁹

Currently, surgery is more often considered for young, active individuals or fractures with high displacement. Indications typically include: (1) open fractures, (2) displaced or comminuted fractures with skin tenting, (3) shortening >2 cm, (4) neurovascular compromise, and (5) associated scapular neck fractures (floating shoulder)[10, 11]. Among operative techniques, open reduction and internal fixation (ORIF) with plates and screws is the most widely preferred, while titanium elastic intramedullary nailing remains an alternative in select cases.¹²

Malunion and shortening may alter scapular kinematics, potentially leading to long-term functional impairment or

Corresponding Author: Zafer Güneş, zafergne@hotmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

glenohumeral arthritis.¹³⁻¹⁷ Although corrective osteotomy can restore alignment, it remains unclear whether delayed correction yields comparable shoulder endurance to early surgical fixation.¹⁸⁻²⁰

The purpose of this study was to compare the clinical and radiological outcomes of midshaft clavicle fractures treated either surgically or conservatively, with a minimum follow-up of one year. We hypothesized that operative fixation would result in faster union and superior early functional recovery, with similar long-term outcomes compared to conservative treatment.

METHODS

Study Design

This retrospective cohort study used data from our institutional orthopedic trauma registry (January 2013-December 2021). The study was approved by the Ankara Training and Research Hospital Ethics Committee (Date: 28.05.2025, Decision No: 503), which waived the requirement for individual informed consent. All procedures adhered to the Declaration of Helsinki.

Patient Selection

A total of 126 adult patients (18-65 years) with midshaft clavicle fractures were identified; 98 met inclusion criteria.

Inclusion criteria: displaced midshaft fractures (Robinson 2B1-2B2), treated surgically or conservatively, with ≥ 12 months of follow-up.

Exclusion criteria: medial/lateral third fractures; pathologic or open fractures; prior shoulder/clavicle surgery; systemic infection; ipsilateral upper-extremity fractures; intramedullary fixation (TEN); incomplete records (Figure 1).

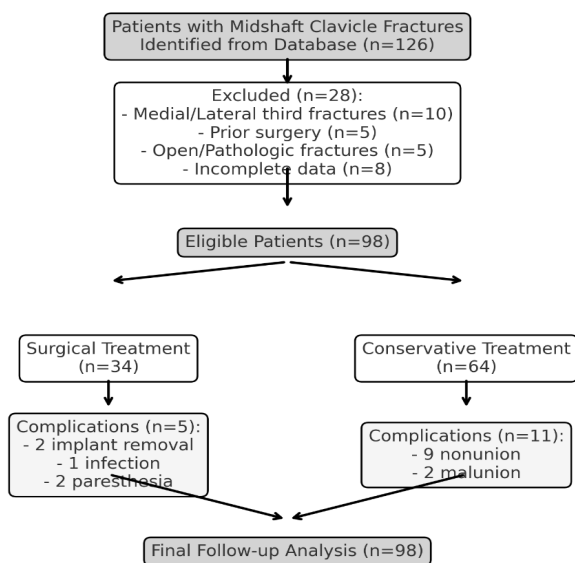


Figure 1. Presents a CONSORT-style patient-selection flowchart

Fracture Classification and Baseline Variables

Fractures were categorized by the Robinson classification. Baseline data included age, sex, fracture type, side, smoking, diabetes, and treatment type.

Treatment Protocols

Operative indications and technique: Surgery was offered for fractures displaced >2 cm, comminuted patterns, skin tenting, neurovascular involvement, or high functional demand.

All procedures were performed by two senior orthopedic surgeons specializing in upper-extremity trauma, with the patient in the beach-chair position under general anesthesia. Superior plating with a pre-contoured 3.5-mm locking plate was used. Prophylactic cefazolin 1 g IV was administered 30 minutes pre-incision.

Passive mobilization began at 2 weeks, with unrestricted motion by 8-12 weeks depending on callus formation.

Conservative treatment: Nonoperative patients were treated with a figure-of-eight brace for six weeks, followed by supervised passive-active range-of-motion exercises under physiotherapy. Full mobilization was permitted after radiologic evidence of callus formation, typically around 12 weeks.

Outcome Measures

Collected variables: age, sex, fracture type, treatment, time to radiologic union, Constant-Murley score,²¹ and Visual Analog Scale (VAS; 0-100 mm)²² for pain at 3, 6, 12 months.

Radiographic union: bridging callus in ≥ 3 cortices on AP and 30° cephalic tilt views + no tenderness/pain on motion.

Nonunion: absence of progressive healing by 6 months±pain or motion at the site.

Malunion: angulation $>15^\circ$ or shortening >20 mm.

Statistical Analysis

Analyses were performed using SPSS v30.0 (IBM, Chicago, USA). Continuous variables were compared with independent-samples T tests; categorical data with Chi-square tests.

Kaplan-Meier survival analysis evaluated time to radiologic union, and log-rank tests assessed intergroup differences. Significance level: $p < 0.05$.

RESULTS

Ninety-eight patients (34 surgical, 64 conservative) were analyzed. Mean ages were 34.6 ± 9.8 and 35.3 ± 10.2 years, respectively. Baseline demographics were comparable (Table 1).

Table 1. Demographic information of patients				
Treatment/characteristic		Surgical, n=34 (34.6%)	Conservative, n=64 (65.4%)	Total, n=98 (100%)
Age (mean)		34.56	35.28	35.03
Gender	Female	15 (44%)	24 (37.5%)	39 (39.7%)
	Male	19 (56%)	40 (62.5%)	59 (60.3%)
Side	Right	21 (61.8%)	41 (64.1%)	62 (63.3%)
	Left	13 (38.2%)	23 (35.9%)	36 (36.7%)
Fracture type	2B1	24 (70.6%)	43 (67.2%)	67 (68.4%)
	2B2	10 (29.4%)	21 (32.8%)	31 (31.6%)

Radiologic union occurred at 16.4 ± 3.2 weeks (surgical) vs 24.3 ± 4.6 weeks (conservative) ($p < 0.001$) (Figure 2, 3).

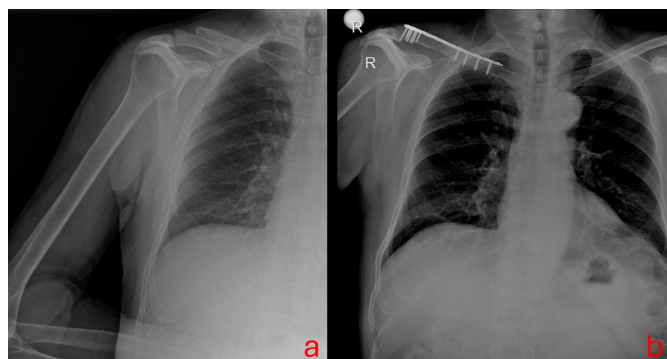


Figure 2. A 38-year-old male patient from the surgical treatment group: (a) preoperative radiographic image, (b) postoperative radiographic at final follow-up

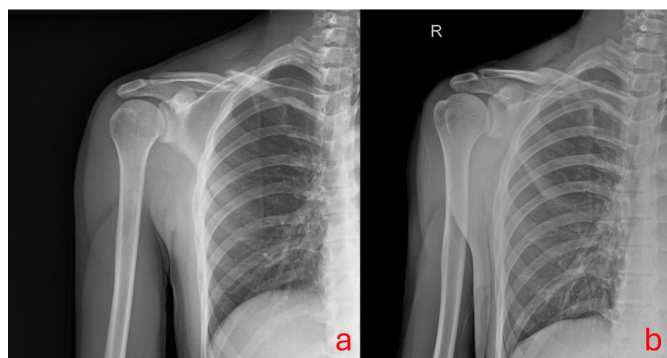


Figure 3. A 32-year-old male patient from the conservative group: (a) radiographic image taken on the day of injury, (b) radiograph at one year

Kaplan-Meier survival analysis confirmed significantly faster union in the surgical group (Log-rank $\chi^2 = 17.46$, $p < 0.001$). Median time-to-union: 15.8 weeks (95% CI 14.9-16.7) for surgical vs 23.6 weeks (95% CI 22.1-25.2) for conservative. At 20 weeks, 88.2% of surgical patients had achieved union versus 39.1% conservative; all achieved complete union by 28 weeks (Figure 4).

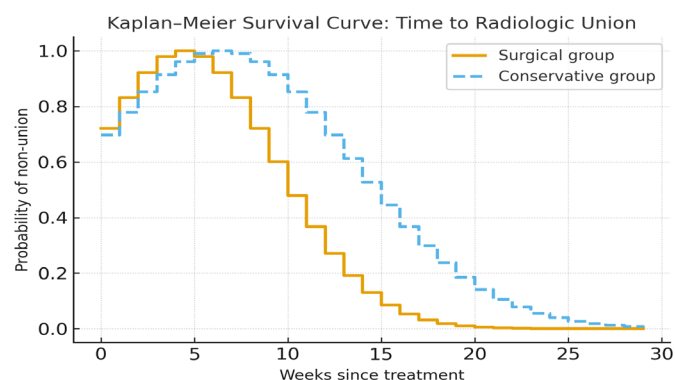


Figure 4. Kaplan-Meier survival analysis

Constant-Murley scores at 3, 6, 12 months:

$79.4 \rightarrow 89.6 \rightarrow 95.7$ (surgical) vs $72.2 \rightarrow 86.8 \rightarrow 95.1$ (conservative). Significant differences at 3 and 6 months ($p < 0.001$), none at 12 months ($p = 0.114$) (Table 2).

VAS pain scores (0-100): $18.8 \rightarrow 12.9 \rightarrow 6.1$ (surgical) vs $22.9 \rightarrow 15.3 \rightarrow 6.6$ (conservative); early differences ($p < 0.001$), none at 12 months ($p = 0.139$) (Table 2).

Table 2. Comparison of mean union time and constant/VAS scores at 3rd, 6th, and 12th months between surgically and conservatively treated patients

Treatment/characteristic	Surgical n=34 (34.6%)	Conservative n=64 (65.4%)	p-value
Union time (week)	16.4	24.3	$p < 0.001$
Constant score (3 rd month)	79.4	72.2	$p < 0.001$
Constant score (6 th month)	89.6	86.8	$p < 0.001$
Constant score (12 th month)	95.7	95.1	0.114
VAS score (3 rd month)	18.8	22.9	$p < 0.001$
VAS score (6 th month)	12.9	15.3	$p < 0.001$
VAS score (12 th month)	6.1	6.6	0.139

VAS: Visual Analog Scale

Complications: 5 (14.7%) surgical vs 11 (17.2%) conservative ($p = 0.260$).

Surgical-2 implant removals, 1 deep infection, 2 transient paresthesias.

Conservative-9 nonunions + 2 malunions, all united after secondary plating.

DISCUSSION

The most important finding of this study was that operative fixation of displaced midshaft clavicle fractures achieved significantly faster radiologic union and superior early functional outcomes, while both treatments yielded similar long-term results once union was complete.

These findings align with major comparative trials. The Canadian Orthopaedic Trauma Society (COTS) found nonunion rates of 15% after nonoperative care vs 2% after plate fixation, with earlier return to function in the surgical cohort.²³ Han et al.²⁴ similarly reported 13% nonunion in conservative treatment and none following surgery. Our series parallels those outcomes, with 0% nonunion after surgery and 14% after nonoperative care.

Biomechanically, the midshaft is a transition zone of bending and torsional stress; loss of length or alignment disrupts scapular motion and may lead to scapulothoracic dyskinesia.²⁵ Clavicular shortening >15 -20 mm measurably impairs elevation and rotation.^{14,15,26} Restoration of anatomic length and stability via rigid fixation thus offers mechanical and biological benefits.²⁷

The superior early Constant and VAS scores in our surgical group reflect these advantages and the ability to mobilize earlier. Woltz et al.²⁸ observed a similar pattern-clear early benefit, convergence by 12 months. The Kaplan-Meier analysis in our study reinforces this temporal advantage, showing a median union nearly eight weeks earlier for surgery.

Surgical intervention, however, carries risks. Implant irritation, cosmetic issues, and infection rates of 10-20% have been documented.²⁹ Our observed complication spectrum was mild; all patients ultimately achieved union.

Different rehabilitation timelines also influence short-term results: early mobilization after plating versus six-week immobilization with a brace delays functional recovery in conservative patients. Thus, functional differences partly reflect protocol design.

While our union criteria combined clinical and radiologic endpoints consistent with practice standards, lack of blinded assessment may have introduced observer bias. Future research should employ CT-based quantitative callus assessment and blinded review.

Potential confounding by indication remains a concern in retrospective designs. Even though we limited inclusion to Robinson 2B fractures, surgical cases may inherently involve greater displacement. Prospective randomized or propensity-matched analyses would strengthen causal inference.

Despite these limitations, our data support that operative fixation yields measurable short-term benefits-faster union and earlier function-while long-term outcomes equalize after healing. In young, high-demand patients, surgery facilitates earlier return to work or sport, whereas conservative management remains appropriate for low-demand, minimally displaced fractures.

In summary, operative fixation restores anatomy and function more rapidly, but ultimate outcome depends on achieving solid union by any method. Treatment should therefore be individualized, balancing early recovery benefits against surgical risk and patient expectations.

Limitations

This study's limitations include its retrospective design and moderate sample size, introducing potential selection bias. Treatment allocation was not randomized, raising the possibility of confounding by indication. Radiographic evaluation was unblinded, follow-up limited to 12 months, and subtle differences in fracture morphology may have persisted despite restricting inclusion to Robinson 2B1-2B2 types.

CONCLUSION

Operative fixation of displaced midshaft clavicle fractures results in significantly faster union and superior early functional recovery, with comparable long-term outcomes once complete healing occurs. Surgical treatment should be favored for young, active individuals or markedly displaced/comminuted fractures, while conservative management remains a reliable alternative for low-demand, minimally displaced cases. Ultimately, treatment should be guided by fracture characteristics, patient activity, and functional expectations.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Ankara Training and Research Hospital Ethics Committee (Date: 28.05.2025, Decision No: 503).

Informed Consent

Because the study was designed retrospectively, no written informed consent form was obtained from patients.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

Data Availability Statement

Data that underlie the conclusions of this research are accessible from the corresponding author upon a reasonable request.

REFERENCES

1. Lenza M, Faloppa F. Surgical interventions for treating acute fractures or non-union of the middle third of the clavicle. *Cochrane Database Syst Rev*. 2015;2015(5):CD007428. doi:10.1002/14651858.CD007428.pub3
2. Robinson CM. Fractures of the clavicle in the adult. Epidemiology and classification. *J Bone Joint Surg Br*. 1998;80(3):476-484. doi:10.1302/0301-620x.80b3.8079
3. Bandovic I, Holme MR, Black AC, Futterman B. Anatomy, Bone Markings. In: StatPearls. Treasure Island (FL): StatPearls Publishing; May 6, 2024.
4. Oki S, Matsumura N, Kiriya Y, Iwamoto T, Sato K, Nagura T. Three-dimensional deformities of nonoperative midshaft clavicle fractures: a surface matching analysis. *J Orthop Trauma*. 2017;31(11):e385-e389. doi:10.1097/BOT.0000000000000931
5. Hoogervorst P, van Schie P, van den Bekerom MP. Midshaft clavicle fractures: current concepts. *EFORT Open Rev*. 2018;3(6):374-380. doi:10.1302/2058-5241.3.170033
6. Frima H, van Heijl M, Michelitsch C, et al. Clavicle fractures in adults; current concepts. *Eur J Trauma Emerg Surg*. 2020;46(3):519-529. doi:10.1007/s00068-019-01122-4
7. Nowak J, Holgersson M, Larsson S. Sequelae from clavicular fractures are common: a prospective study of 222 patients. *Acta Orthop*. 2005; 76(4):496-502. doi:10.1080/17453670510041475
8. Naveen BM, Joshi GR, Hari Krishnan B. Management of mid-shaft clavicular fractures: comparison between non-operative treatment and plate fixation in 60 patients. *Strategies Trauma Limb Reconstr*. 2017;12(1): 11-18. doi:10.1007/s11751-016-0272-4
9. Zlowodzki M, Zelle BA, Cole PA, Jeray K, McKee MD; Evidence-Based Orthopaedic Trauma Working Group. Treatment of acute midshaft clavicle fractures: systematic review of 2144 fractures: on behalf of the evidence-based orthopaedic trauma working group. *J Orthop Trauma*. 2005;19(7):504-507. doi:10.1097/01.bot.0000172287.44278.ef
10. Song HS, Kim H. Current concepts in the treatment of midshaft clavicle fractures in adults. *Clin Shoulder Elb*. 2021;24(3):189-198. doi:10.5397/cise.2021.00388
11. Biz C, Pozzuoli A, Belluzzi E, et al. An institutional standardised protocol for the treatment of acute displaced midshaft clavicle fractures (ADMCFs): conservative or surgical management for active patients? *Healthcare (Basel)*. 2023;11(13):1883. doi:10.3390/healthcare11131883
12. Li L, Yang X, Xing F, Jiang J, Tang X. Plate fixation versus intramedullary nail or Knowles pin fixation for displaced midshaft clavicle fractures: a meta-analysis of randomized controlled trials. *Medicine (Baltimore)*. 2020;99(39):e22284. doi:10.1097/MD.00000000000022284
13. Hung LW, Lu HY, Chen TY, Wang TM, Lu TW. Residual kinematic deviations of the shoulder during humeral elevation after conservative treatment for mid-shaft clavicle fractures. *Front Bioeng Biotechnol*. 2024;12:1413679. doi:10.3389/fbioe.2024.1413679
14. Matsumura N, Ikegami H, Nakamichi N, et al. Effect of shortening deformity of the clavicle on scapular kinematics: a cadaveric study. *Am J Sports Med*. 2010;38(5):1000-1006. doi:10.1177/0363546509355143

15. Hillen RJ, Burger BJ, Pöll RG, van Dijk CN, Veeger DH. The effect of experimental shortening of the clavicle on shoulder kinematics. *Clin Biomech (Bristol)*. 2012;27(8):777-781. doi:10.1016/j.clinbiomech.2012.05.003
16. Weinberg DS, Vallier HA, Gaumer GA, Cooperman DR, Liu RW. Clavicle fractures are associated with arthritis of the glenohumeral joint in a large osteological collection. *J Orthop Trauma*. 2016;30(11):605-611. doi:10.1097/BOT.0000000000000654
17. Pailhes RG, Bonnevalle N, Laffosse J, Tricoire J, Cavaignac E, Chiron P. Floating shoulders: clinical and radiographic analysis at a mean follow-up of 11 years. *Int J Shoulder Surg*. 2013;7(2):59-64. doi:10.4103/0973-6042.114230
18. Potter JM, Jones C, Wild LM, Schemitsch EH, McKee MD. Does delay matter? The restoration of objectively measured shoulder strength and patient-oriented outcome after immediate fixation versus delayed reconstruction of displaced midshaft fractures of the clavicle. *J Shoulder Elbow Surg*. 2007;16(5):514-518. doi:10.1016/j.jse.2007.01.001
19. Kluijfhout WP, Tutuhaturnewa ED, van Olden GDJ. Plate fixation of clavicle fractures: comparison between early and delayed surgery. *J Shoulder Elbow Surg*. 2020;29(2):266-272. doi:10.1016/j.jse.2019.06.022
20. Sanchez PH, Fleury IG, Parker EA, et al. Early versus delayed surgery for midshaft clavicle fractures: a systematic review. *Iowa Orthop J*. 2023;43(1):151-160.
21. Constant CR, Gerber C, Emery RJ, Søjbjerg JO, Gohlke F, Boileau P. A review of the Constant score: modifications and guidelines for its use. *J Shoulder Elbow Surg*. 2008;17(2):355-361. doi:10.1016/j.jse.2007.06.022
22. Delgado DA, Lambert BS, Boutris N, et al. Validation of digital visual analog scale pain scoring with a traditional paper-based Visual Analog Scale in adults. *J Am Acad Orthop Surg Glob Res Rev*. 2018;2(3):e088. doi:10.5435/JAOSGlobal-D-17-00088
23. Canadian Orthopaedic Trauma Society. Nonoperative treatment compared with plate fixation of displaced midshaft clavicular fractures. A multicenter, randomized clinical trial. *J Bone Joint Surg Am*. 2007;89(1):1-10. doi:10.2106/JBJS.F.00020
24. Han X, Zhang Y, Zhang X, Tan J. Comprehensive comparison between conservative therapy and surgical management for completely displaced and comminuted mid-shaft clavicle fractures. *Int Orthop*. 2024;48(7):1871-1877. doi:10.1007/s00264-024-06198-1
25. Shields E, Behrend C, Beiswenger T, et al. Scapular dyskinesis following displaced fractures of the middle clavicle. *J Shoulder Elbow Surg*. 2015;24(12):e331-e336. doi:10.1016/j.jse.2015.05.047
26. Kim D, Lee D, Jang Y, Yeom J, Banks SA. Effects of short malunion of the clavicle on in vivo scapular kinematics. *J Shoulder Elbow Surg*. 2017;26(9):e286-e292. doi:10.1016/j.jse.2017.03.013
27. Hung LW, Lu HY, Chang CH, Chen TY, Wang TM, Lu TW. Effects of internal fixation for mid-shaft clavicle fractures on shoulder kinematics during humeral elevations. *Front Bioeng Biotechnol*. 2021;9:710787. doi:10.3389/fbioe.2021.710787
28. Woltz S, Krijnen P, Schipper IB. Plate fixation versus nonoperative treatment for displaced midshaft clavicular fractures: a meta-analysis of randomized controlled trials. *J Bone Joint Surg Am*. 2017;99(12):1051-1057. doi:10.2106/JBJS.16.01068
29. Wijdicks FJ, Van der Meijden OA, Millett PJ, Verleisdonk EJ, Houwert RM. Systematic review of the complications of plate fixation of clavicle fractures. *Arch Orthop Trauma Surg*. 2012;132(5):617-625. doi:10.1007/s00402-011-1456-5

Determination of reproductive health attitudes and affecting factors of women aged 15-49

✉Nevin Utkualp¹, ✉Melih Başoğlu²

¹Department of Obstetrics and Gynecology Nursing, Faculty of Health Sciences, Bursa Uludağ University, Bursa, Türkiye

²Department of Obstetrics and Gynecology Nursing, Faculty of Health Sciences, Dokuz Eylül University, İzmir, Türkiye

Cite this article as: Utkualp N, Başoğlu M. Determination of reproductive health attitudes and affecting factors of women aged 15-49. *J Health Sci Med.* 2025;8(6):1176-1181.

Received: 29.07.2025

Accepted: 21.10.2025

Published: 25.10.2025

ABSTRACT

Aims: This study aimed to determine women's attitudes towards reproductive health and the factors affecting them.

Methods: The study was a descriptive and cross-sectional study. The study population consisted of married women of reproductive age (15-49 years) who applied to the Bursa Uludağ University Health Application Research Center Gynecology and Obstetrics Outpatient Clinic. The Descriptive Information Form and Determination of Married Women's Reproductive Health Protective Attitudes Scale (RHPAS) were used to collect data from the women. The research data were collected by face-to-face interviews between February and July 2020. The SPSS v26 package program was used to evaluate the data, and the significance test of the difference between two means in independent groups (T test) was used.

Results: The mean age of the women was found to be 34.35±9.17 years. 36.8% of the women were high school graduates, and 53.4% were housewives. While 56.9% of the women did not perform breast self-examination (BSE), 89.1% did not know when to perform BSE. 54% of the women applied to the gynecology and obstetrics outpatient clinic for ultrasonography (USG), 48.9% for pap smear, 28.2% for biopsy, and 36.2% for hormone tests. The mean score of the Determination of Women's RHPAS was 148.28±17.38 (min: 104, max: 189). Women who reported having regular gynecological examinations had higher mean scores on the RHPAS total score. Women who exhibited the behavior of performing BSE had higher mean scores on the RHPAS total score in the study.

Conclusion: Opportunistic information programs can be managed at different health levels, both in rural and urban areas, so that women of reproductive age can develop attitudes that protect reproductive health. Public service announcements can be created to provide mass access and raise awareness at the social level, and people can be invited to health services to receive SRH education.

Keywords: Women, reproductive health, protective attitude

INTRODUCTION

Reproductive health is an issue that directly affects women's lives. The World Health Organization (WHO) has defined reproductive health as "not only the absence of disease and disability related to reproductive system functions and processes, but also a state of complete physical, mental, and social well-being related to all these".^{1,2} Reproductive health also includes having the ability to reproduce and experiencing the freedom to make decisions in using reproductive ability.^{3,4} The acquisition of positive health behaviors by women, who assume important roles, especially in the health management of the family, is necessary for a healthy society.⁵ The protection of women's health is not only important for them directly, but also means the protection of family and community health.⁶ Continuous changes in social structure, developments in science and technology make the need for education compulsory.⁷ In our country, reproductive health is a concept that concerns both women and men, but women are more

likely to have problems. This is because women go through periods such as menstruation, pregnancy, childbirth, and puerperium, which can lead to reproductive health problems in women. Health awareness means being aware of how to protect oneself from diseases, what to do to be healthier, and what to do to maintain it.⁸ Improving the sexual and reproductive health of women of reproductive age is also one of the main goals of the Sustainable Development Goals.^{5,6} It contributes to maternal and child health by supporting the development of adequate and accurate information, positive reproductive health, and regular health screenings.^{3,4} Health policies carried out to protect the well-being of individuals are extremely important for the socio-economic development of the country in the coming years.¹ However, socio-cultural norms and taboos against women in societies limit access to sexual and reproductive health services.⁶ Furthermore, socially constructed roles and responsibilities, and the perception

Corresponding Author: Nevin Utkualp, nutkualp@uludag.edu.tr



This work is licensed under a Creative Commons Attribution 4.0 International License.

that services such as care and treatment will be costly, place women in a disadvantaged position in terms of health status and access to health services.^{6,7} Currently, socio-economic conditions and increasing stress have a negative impact on reproductive health outcomes. Some of these effects on women range from increased risk of disease to malnutrition, lack of access to reproductive health services, adverse obstetric outcomes, and even death.⁹ Nurses who are knowledgeable and capable of counseling are needed to solve these problems. Nurses should recognize the risks of married women to whom they provide care and should be able to provide appropriate counseling to the individual when necessary.¹⁰ Nurses are able to make important contributions to raising awareness by informing women about reproductive health and thus reducing the negative effects on reproductive health.¹¹ This study was conducted to determine women's attitudes towards reproductive health and the factors affecting them.

METHODS

Research Ethics

The necessary institutional permission was obtained from Bursa Uludağ University Health Practice and Research Center Department of Obstetrics and Gynecology. Non-interventional research ethics committee approval was obtained at Bursa Uludağ University Faculty of Medicine Clinical Researches Ethics Committee (Date: 01.06.2020, Decision No: 2020-03). All procedures were carried out in accordance with the ethical rules and the principles of the Declaration of Helsinki.

Type of Research

The research is a descriptive and cross-sectional study.

Research Population and Sample

The population of the study consisted of married women of reproductive age (15-49 years) who applied to Bursa Uludağ University Health Application Research Center Gynecology and Obstetrics Outpatient Clinic. The study was conducted randomly. The inclusion criteria were volunteering to participate in the study and being of reproductive age. Exclusion criteria were not speaking Turkish, being illiterate, being single, and being in menopause. The required effect size to be used in the sample calculation of the study was calculated through the research conducted by Kılınç¹¹ using the same scale. Accordingly, with an effect size of 0.5, an alpha value of 0.05 (type I error), and a power of 0.80, the minimum sample size was determined as 102. 174 women were reached within the scope of this research.

Data Collection Tools

The Descriptive Information Form and Determination of Married Women's Reproductive Health Protective Attitudes Scale (RHPAS) were used to collect data from the women. The Descriptive Information Form is a 32-question form prepared by researchers and includes questions about women's sociodemographic characteristics, gynecological and obstetric history, attitudes, and behaviors towards sexual and reproductive health. Women's protective attitudes and behaviors towards reproductive health were evaluated by

means of the 39-question RHPAS developed by Demirci.³ In the five-point Likert-type scale, the questions are scored as "1=Never, 2=Rarely, 3=Sometimes, 4=Most of the time, and 5=Always". Questions 5, 10, 16, and 28 are reverse-scored. In reverse questions, the answer was re-scored as 5 if it was 1, 1 if it was 5, 4 if it was 2, 2 if it was 4, and 3 was accepted as the same. The minimum score that can be obtained from the scale is 39, and the maximum score is 195. The scale includes; behavior of visiting a doctor in matters concerning reproductive health (13th, 14th, 15th, 16th, 24th, 25th, 33rd and 39th questions), prevention of reproductive organ and breast cancer (20th, 21st, 22nd and 23rd questions), general health behaviors for protection of reproductive health (1st, 29th, 30th, 31st, 32nd, 33rd, 34th, 35th, 36th, 37th and 38th questions) protection from genital tract infections (2nd, 3rd, 4th, 5th, 6th, 7th, 8th, 9th, 10th, 11th, 12th, 17th, 18th and 19th questions) and protection from unintended pregnancy (26th, 27th and 28th questions). The Cronbach-alpha coefficient of the scale is 0.82. In this study, the Cronbach-alpha coefficient was 0.80. Verbal and written permission were obtained from the author for the use of the scale.

Data Collection

The purpose of the study was explained to the participants, and their written and verbal consent was obtained. The research data were collected by face-to-face interviews between February and July 2020. Forms were distributed to women who applied to the Bursa Uludağ University Health Application Research Center Gynecology and Obstetrics Outpatient Clinic and who agreed to participate in the study. The women were given an average of 20 minutes to complete the forms.

Statistical Analysis

The Statistical Package for the Social Sciences (SPSS) v26 package program was used to evaluate the data. The data were evaluated with the significance test of the difference between two means in independent groups (T test). The data collected in the study were expressed by number, percentage, arithmetic mean, and standard deviation from descriptive statistical methods. The level of statistical significance in the study was accepted as $p < 0.05$.

RESULTS

Sociodemographic Characteristics of Women

The mean age of the women was found to be 34.35 ± 9.17 years. 92.5% of the women had social security, and 74.5% of those with social security had Social Insurance Institution (SSK) insurance. 36.8% of the women were high school graduates, and 53.4% were housewives. 78.7% of women had a nuclear family type, and 67.2% of women stated their income perception as "my income is equal to my expenses". 75.9% of the women did not have any harmful habits (smoking, alcohol, drugs), and 19.5% were smokers. The sociodemographic characteristics of the women are given in [Table 1](#).

Obstetric and Gynecological Characteristics of Women

The mean number of pregnancies of women was 1.50 ± 1.45 , and their mean number of children was 1.12 ± 1.11 . The

Table 1. Sociodemographic characteristics of women

	Mean±SD	Min-max
Age	34.35 ± 9.17	17-49
	Number (n)	Percent (%)
Social security		
Retirement fund	21	13.0
SSI	120	74.5
SSO for artisans and the self-employed	14	8.7
Private health insurance	6	3.7
Level of education		
Literate (primary school)	22	12.6
Secondary education	26	14.9
High school	64	36.8
College and upper	62	35.6
Occupation		
Housewife	93	53.4
Officer	22	12.6
Independent business	6	3.4
Other	53	30.5
Family type		
Nuclear family	137	78.7
Extended family	32	18.4
Broken family	5	2.9
Income status perception		
Income less than expense	51	29.3
Income equal to expense	117	67.2
Income more than expense	6	3.4
Bad habits		
Smoking	34	19.5
Alcohol	8	4.6
No bad habits	132	75.9

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

mean age of their menarche was 13.09±1.61 years. Among the women, 42.5% reported irregular menstrual cycles and 43.1% reported sometimes dysmenorrhea. While 56.9% of the women did not perform breast self-examination (BSE), 89.1% did not know when to perform BSE. 54% of the women applied to the gynecology and obstetrics outpatient clinic for ultrasonography (USG), 48.9% for pap smear, 28.2% for biopsy, and 36.2% for hormone tests. 25.3% of the women had a history of curettage, and 19.5% had a history of abortion. 71.8% of the women stated that they did not use any contraceptive method. Furthermore, 48.3% of the women reported having gynecological examinations, and 64.4% of the women reported that they applied to health services for health checks only when they had a health problem. Regarding the genital hygiene habits of the women, it was observed that 77% of the women performed genital cleansing with water only, 81.6% did it from front to back, and 78.2% changed their underwear every day. 91.4% of the women reported that they preferred to wear cotton underwear. Regarding the menstrual hygiene habits of the women, 27.6% of the women used pads daily, 63.2% changed pads every four hours during menstruation with little bleeding, and 69% changed pads every one- or two-hours during menstruation with heavy bleeding. It was also found that 85.6% of the women washed their hands before using the pad, and 77% of them paid attention not to handle the part of the pad touching the vagina while placing the pad. 53.4% of the women had the habit of vaginal douching, and 83.9% reported that they did not avoid taking a shower during

menstruation. The obstetric and gynecological characteristics of the women are presented in [Table 2](#).

Table 2. Obstetric and gynecological characteristics of women

	Mean±SD*	Min-max
Number of pregnancies	1.50±1.45	1-5
Number of children	1.12 ± 1.11	0-5
Menarche age	13.09 ± 1.61	9-18
	Number (n)	Percent (%)
Irregularity in menstrual cycle		
Yes	74	42.5
No	100	57.5
Dysmenorrhea		
Yes	43	24.7
Sometimes	75	43.1
No	56	32.2
Regular BSE* behavior		
Yes	75	43.1
No	99	56.9
BSE method		
She knows the right method	19	10.9
She knows the wrong method	155	89.1
Gynecological examinations applied**		
Ultrasonography	94	54.0
Pap smear	85	48.9
Biopsy	49	28.2
Hormonal examination	63	36.2
Curettage history		
Yes	44	25.3
No	130	74.7
Abortus history		
Yes	34	19.5
No	140	80.5
Contraceptive method		
She uses	49	28.2
She does not use	125	71.8
Having regular gynaecological examination		
She does	84	48.3
She does not	90	51.7
The status of having a health check		
She sees a doctor when she has a complaint	112	64.4
She sees a doctor regularly, even though she has no complaints	46	26.4
She does not see a doctor at all	16	9.2
The method of cleaning the genital area		
Only using water	134	77.0
Using water and soap	40	23.0
The direction of cleaning the genital area		
Front to back	142	81.6
Back to front	32	18.4
The frequency of changing underwear		
Everyday	136	78.2
Every other day	38	21.8
The behavior of vaginal douching		
Yes	93	53.4
No	81	46.6
The behavior of washing hands following pad changing		
Yes	149	85.6
No	25	14.4
Bathing during menstruation		
Yes	146	83.9
No	28	16.1

*BSE: Breast self-examination, **The participants selected multiple options, SD: Standard deviation, Min: Minimum, Max: Maximum

Findings Regarding Women's Attitudes Towards Protecting Reproductive Health

In the study, the mean score of the Determining Reproductive Health Protective Attitudes Scale (RHPAS) was 148.28 ± 17.38 (min: 104, max: 189). The mean scores of the women in the sub-dimensions of the scale were found to be 31.18 ± 6.13 for the sub-dimension of the behavior of going to the doctor for issues related to reproductive health, 10.20 ± 4.38 for the sub-dimension of protection from reproductive organ and breast cancer, 34.79 ± 6.04 for the sub-dimension of general health behaviors for reproductive health protection, 60.97 ± 5.99 for the sub-dimension of protection from genital tract infections and 11.11 ± 3.05 for the sub-dimension of protection from unwanted pregnancies. Table 3 shows the mean scores of the women in the sub-dimensions and total scores of the RHPAS.

	Mean $\pm$ SD	Min-max
The behavior of seeing a doctor concerning reproductive health	31.18 ± 6.13	16-40
Protection against reproduction, organ, and breast cancer	10.20 ± 4.38	4-20
General health behaviors for protecting reproductive health	34.79 ± 6.04	21-50
Protection against genital tract infections	60.97 ± 5.99	32-97
Protection against unintended pregnancy	11.11 ± 3.05	3-15
RHPAS total score	148.28 ± 17.38	104-189
RHPAS: Reproductive Health Protective Attitudes Scale, SD: Standard deviation, Min: Minimum, Max: Maximum		

A statistically significant difference was found between the women's behaviors of performing BSE and the mean total score of the RHPAS ($p < 0.05$). It was found that the mean total RHPAS score of the women who exhibited the behavior of performing BSE was higher. A statistically significant difference was found between the women's USG status and the mean total score of the RHPAS ($p < 0.05$). The mean total RHPAS score of the women who had undergone USG was higher. A statistically significant difference was found between the women's pap smear status and their mean total score of the RHPAS ($p < 0.05$). It was observed that the mean total RHPAS score of the women who had a Pap smear was higher. A statistically significant difference was found between the women's biopsy status and the mean total score of the RHPAS ($p < 0.05$). The RHPAS total mean scores of women undergoing biopsy are higher. A statistically significant difference was found between the women's contraceptive method use status and the mean total score of the RHPAS ($p < 0.001$). It was found that women who reported using contraceptive methods had higher mean total scores on the RHPAS. A statistically significant difference was found between the women's status of having regular gynecologic examination and the mean total score of the RHPAS ($p < 0.05$). Women who reported having regular gynecological examinations were found to have higher mean RHPAS scores. The mean total scores of the women according to some characteristics are given in Table 4.

DISCUSSION

Women's protective attitudes towards reproductive health were found to be slightly above average in our study. In the study

Table 4. RHPAS total mean scores according to some characteristics of women

RHPAS total score, mean±SD*		p**
Irregularity in menstrual cycle		
Yes	146.24±15.82	.18
No	149.79±18.39	
Regular BSE behavior		
Yes	152.82±17.72	.002**
No	144.83±16.39	
Ultrasonography		
She does	151.59±18.32	.005**
She does not	144.38±15.44	
Pap smear		
She does	152.54±17.95	.001**
She does not	144.21±15.88	
Biopsy		
She does	154.53±17.44	.003**
She does not	145.83±16.81	
Hormonal examination		
She does	150.09±18.76	.301
She does not	147.25±16.55	
Curretage history		
Yes	150.29±20.10	.376
No	147.60±16.40	
Abortus history		
Yes	151.67±18.91	.205
No	147.45±16.96	
Contraceptive method		
She uses	155.73±16.34	.000**
She does not use	145.36±16.96	
Having regular gynaecological examination		
She does	152.27±17.05	.003**
She does not	144.55±16.95	
RHPAS: Reproductive Health Protective Attitudes Scale, *SD: Standard deviation, **p<0.05, BSE: Breast self-examination		

conducted by Erkin and Şahin, the mean score for "general health behaviors aimed at protecting reproductive health" was found to be similar to the results of our study.¹⁰ Another study carried out in our country emphasized that the high average score for reproductive health was due to women receiving education based on the healthy lifestyle model.¹¹ Half of the women reported that they had gynecological examinations, and slightly more than half of the women reported that they applied to health services for health check-ups only when they had a health problem.¹² It is expected that women should have reproductive health protective attitudes in the promotion of reproductive health. In our study, it was determined that more than half of the women did vaginal douching, and the majority avoided douching during menstruation. It is thought that these results are due to the fact that one-third of the participants are university graduates, and more than half of them are housewives; they do not have sufficient knowledge,

and their access to information is limited. The studies performed are similar to our study.¹³ Unlike our results, in another study conducted in China, it was reported that the reproductive health knowledge score of more than half of the reproductive health women was associated with a low education level.¹⁴ It has been reported that Tibetan women living in China know about reproductive health, but their attitudes and behaviors are inadequate.¹⁵ In a study of Filipino women workers working in the United States, it was reported that culture and religion are an obstacle in seeking care where reproductive health interests are inadequate.¹⁶ Women in Saudi Arabia have reported socio-economic and cultural limitations to the formation of awareness.² In our study, it was determined that it increased reproductive health attitudes in all sub-dimensions and the total. In this context, women's attitudes towards protecting reproductive health are critical in the development of expected reproductive health behaviors. It also enables women to use preventive health services correctly, develop behaviors to improve reproductive health, and raise conscious awareness for preventable maternal and infant mortality. In our study, the reproductive health averages of women who exhibited the behavior of performing BSE were positively affected. In the study conducted by Maurren et al.,¹⁷ which is similar to our study, women's knowledge of the signs and symptoms of breast cancer and their likelihood of performing BSE increased. Therefore, it was reported that it increased reproductive health attitudes in a study in which women were given training on reproductive health, and the results were evaluated.¹³ In Kara's¹³ study, it was determined that, in the presence of other variables that may affect attitudes toward protecting reproductive health, women's behaviors related to physical activity, health responsibility, and spiritual development were associated with their attitudes toward maintaining reproductive health. Kazami¹⁸ emphasized that paying attention to reproductive health at every stage of a woman's life is of vital importance for her overall health. In Cho's¹⁹ study, it was reported that adolescent women tend to conceal reproductive health problems. In Aydoğdu's²⁰ study, it was determined that after the training given, the frequency of performing BSE, mammography, and pap smear tests increased in the experimental group compared to the control group. The results obtained in Kılınç's⁵ study are similar to the results of other studies in which the effect of training on reproductive health cancers on women's knowledge, attitudes, and behaviors is revealed. In a study conducted by Davey-Rothwell et al.²¹ in the United States of America, where HIV and CYBH rates are increasing despite numerous behavioral interventions, it was determined that the education and three-month follow-up decreased risky sexual behaviors. It is thought that women may have cared more about going to the doctor when awareness was raised about the causes of these symptoms and the importance of early diagnosis. Furthermore, contraceptive method use was found to have a positive effect on reproductive health behaviors in the study results. In Dhakai's²² study, it was reported that women had knowledge but were inadequate in practice. In another study similar to our study, in a systematic review of 59 studies from 22 different countries, it was found that Muslim women have personal, religious, cultural, or spousal barriers to access to

sexual and reproductive health services and education, and therefore, women's knowledge levels are insufficient and access to reproductive health services is limited.² In another study on the subject, it was pointed out that there are social and cultural reasons why they get information about contraceptive pregnancy from the internet, but this accessible information should be healthy and reliable.²³ For these reasons, regular training activities of nurses will help to protect reproductive health. In line with these results, it is considered important for women to receive counseling from healthcare professionals in order to promote behaviors that support reproductive health.

CONCLUSION

As a result of the study, it was found that women's reproductive health protective attitudes were at the middle level, and other subscale scores were slightly better. It is thought that interventions that improve women's healthy lifestyle behaviors should be continued, women should be followed up for a longer period of time to change their attitudes, and their spouses should also participate in the training to achieve the targeted women's health profile. Access to reproductive health services should be improved, and nurses working in these centers have great responsibilities concerning the matter. Nurses can raise public reproductive health awareness by ensuring that women are conscious and sensitive about their health and make the right decisions. Especially considering the use of social media in today's world, they can make calls to raise awareness on reproductive health through social media tools to reach women.

ETHICAL DECLARATIONS

Ethics Committee Approval

The study was carried out with the permission of the Bursa Uludağ University Faculty of Medicine Clinical Researches Ethics Committee (Date: 01.07.2020, Decision No:2020-03).

Informed Consent

All patients signed and free and informed consent form.

Referee Evaluation Process

Externally peer-reviewed.

Conflict of Interest Statement

The authors have no conflicts of interest to declare.

Financial Disclosure

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

Author Contributions

All of the authors declare that they have all participated in the design, execution, and analysis of the paper, and that they have approved the final version.

REFERENCES

1. Health Organization (WHO). Reproductive health. Erişim adresi <https://www.who.int/nursing-midwifery-resources/en/> - 78k.pdf. 13 Eylül 2025.

2. Alomair N, Alageel S, Davies N, Bailey JV. Factors influencing sexual and reproductive health of muslim women: a systematic review. *Reprod Health*. 2020;17(1):33. doi:10.1186/s12978-020-0888-1
3. Demirci H. Evli Kadınların Üreme Sağlığını Koruyucu Tutumlarını Belirleme Ölçeği'nin geliştirilmesi. 4. Uluslararası Üreme Sağlığı ve Aile Planlaması Kongresi, 2005:243. Manisa.
4. Demirbağ CB, Çan G, Köksal İ, Kaya S. Üreme sağlığını koruma ve geliştirmede farklı yaklaşım: anne eğitiminin kızlarına aktarılmasının değerlendirilmesi. *Electr J Vocat Colleg*. 2013;3(1):136-148.
5. Kılıç NÖ. Sağlığı Geliştirme Modeline Göre Verilen Eğitimin Kadınların Üreme Sağlığı Tutumlarına ve Öz-yeterliliklerine Etkisi (Yayımlanmamış doktora tezi) Atatürk Üniversitesi Sağlık Bilimleri Enstitüsü Halk Sağlığı Hemşireliği Anabilim Dalı.Erzurum. 2015.
6. İlçioğlu K, Keser N, Çınar N. Women's health and affecting factors in Turkey. *J Human Rhythm*. 2017;3(3):112-119.
7. Aslan E, Bektaş H, Başgöl Ş, Demir S, Vural P. Knowledge and behaviour of university students related to sexual health. *STED*. 2014;23(5):174-182.
8. Öncel S, Teskereci G. reproductive health according to life periods in women and nursing approach. *J Public Health Nurs*. 2018;4(1):53-69.
9. ACOG publications Obstetrics & Gynecology October 2018;132(4):1075-1076. Erişim adresi <https://journals.lww.com/greenjournal/toc/2024/10000>
10. Erkin G, Şahin E. The effect of reproductive health protective attitudes of married women on sexual life. *Kırşehir Ahi Evran Univ J Health Sci*. 2023;7(3):214-225.
11. Kılınc NÖ, Hacıaloğlu N. The effect of training given according to the health promotion model on women's reproductive health attitudes and self-efficacy. *Int J Nurs*. 2021;8(2):70-81. doi:10.15640/ijn.v8n2a8
12. Erbil N, Göktaşlar Z. Determination of affecting factors and the reproductive health protective attitudes of married women. *Turk Clin J Gynecol Obst*. 2010;20(5):295-303.
13. KaraP. The relationship between healthy lifestyle behaviors and reproductive health protective attitudes of women in reproductive. *CBU-SBED*. 2024;11(3):371-378.
14. Chen M, Luo Y, Fu J, et al. Reproductive health status and related knowledge among women aged 20-39 years in rural China: a cross-sectional study. *Reprod Health*. 2020;17(1):90-100. doi:10.1186/s12978-020-00939-2
15. Gipson JD, Moucheraud C, Gyaltzen K, Tsering L, Nobari TZ, Gyal L. Nomadic Tibetan women's reproductive health: findings from cross-sectional surveys with a hard-to-reach population. *Reprod Health*. 2021;18(1):63-71. doi:10.1186/s12978-020-01052-0
16. Nagtalon-Ramos J, Ayres C, Faught B. Sexual and reproductive health knowledge, attitudes, and self-efficacy among young adult Filipino American women. *J Transcult Nurs*. 2022;33(5):576-584. doi:10.1177/10436596221107600
17. Koech JM, Magutah K, Mogere DM, et al. Knowledge, attitude and practices around breast cancer and screening services among women of reproductive age in Turbo sub-county, Kenya. *Heliyon*. 2024;10(11):e31597. doi:10.1016/j.heliyon.2024.e31597
18. Fathnezhad-Kazemi A, Zeighami Mohammadi S, Nayeibinia AS. Association between health-promoting behaviors and reproductive health needs in women experiencing domestic violence. *J Interpers Violence*. 2023;38(19-20):10998-11016. doi:10.1177/08862605231178361
19. Cho I, Park YJ. The effectiveness of a tailored programme to promote reproductive-health-promoting behaviour in young women based on the precaution adoption process model: a randomized controlled trial. *Nurs Open*. 2023;10(3):1704-1714. doi:10.1002/nop2.1425
20. Aydoğdu NG (2011). Sağlığı geliştirme modellerine göre yapılan hemşirelik girişimlerinin yoksul kadınların meme ve serviks kanserine yönelik erken tanı davranışlarına etkisi. (Yayımlanmış doktora tezi). Dokuz Eylül Üniversitesi Sağlık Bilimleri Enstitüsü Halk Sağlığı Hemşireliği Anabilim Dalı. İzmir.
21. Davey-Rothwell MA, Tobin K, Yang C, Sun CJ, Latkin CA. Results of a randomized controlled trial of a peer mentor HIV/STI prevention intervention for women over an 18 month follow-up. *AIDS Behav*. 2011;15(8):1654-1663. doi:10.1007/s10461-011-9943-9
22. Dhakal U, Shrestha RB, Bohara SK, Neupane S. Knowledge, attitude and practice on family planning among married muslim women of reproductive age. *J Nepal Health Res Counc*. 2020;18(2):238-242. doi:10.33314/jnhrc.v18i2.2244
23. Timilsina A, Bhandari B, Johns A, Thapa S. Barriers and facilitators to self-care practices for sexual and reproductive health among women of reproductive age. *PLoS One*. 2024;19(5):e0303958. doi:10.1371/journal.pone.0303958

Pregnancy outcomes in women with Behçet's disease

 Erhan Muluk

Department of Obstetrics and Gynecology, Antalya Anatolia Hospital, Antalya, Türkiye

Cite this article as: Muluk E. Pregnancy outcomes in women with Behçet's disease. *J Health Sci Med.* 2025;8(6):1182-1186.

Received: 22.09.2025

Accepted: 21.10.2025

Published: 25.10.2025

ABSTRACT

Aims: To describe maternal and neonatal outcomes of pregnancies complicated by Behçet's disease (BD) based on a pooled analysis of published case reports over the past 30 years.

Methods: A systematic search of PubMed, Web of Science Core Collection, and Google Scholar was performed to identify case reports and case series of BD in pregnancy published between January 1995 and March 2025. Data on maternal age, disease duration, disease activity, delivery outcomes, and neonatal birth weight were extracted.

Results: A total of 61 cases were identified. The mean maternal age was 32.3 ± 6.3 years, and the mean disease duration was 5.2 ± 3.2 years. Flares during pregnancy were reported in 32.8% of women. Preterm delivery (11.5%) and low birth weight (11.5%) were the most common complications, while miscarriage (3.3%) and preeclampsia (1.6%) were less frequent. The mean neonatal birth weight was 3045 ± 517 g.

Conclusion: In this pooled analysis, preterm delivery and low birth weight were the most frequent complications among pregnancies with BD. Obstetricians should be aware that BD may first present during pregnancy and should closely monitor affected women, especially in high-prevalence regions.

Keywords: Behçet's disease, pregnancy, preterm birth, low birth weight, vasculitis

INTRODUCTION

Behçet's disease (BD) is a systemic vasculitis that may affect arteries and veins of all sizes. It is characterized by recurrent oral, genital, and intestinal ulcers, skin lesions, uveitis, and, in some patients, parenchymal brain involvement.¹ Owing to its distinctive geographic distribution, it has also been called "silk road disease." Although rare on a global scale, BD is a chronic, multisystemic inflammatory condition with a complex pathogenesis involving infectious, genetic, epigenetic, and immunological factors.²

The effect of BD on pregnancy remains uncertain. There is no consensus regarding the true risk of maternal or neonatal complications. Some studies suggest increased rates of miscarriage and intrauterine growth restriction, while others report higher rates of cesarean delivery among affected women.^{3,4} Predicting the course of BD during pregnancy is particularly challenging, as the disease may remain quiescent in some patients while flaring in others.⁵

In terms of global distribution, BD is most prevalent in Türkiye and Iran, where prevalence estimates are among the highest worldwide.⁶ According to epidemiological reports spanning 1974 to 2015, Türkiye consistently ranks first, followed by the Middle East, Asia, and Europe.⁷ Obstetricians practicing in these regions, particularly in countries along the Silk Road where BD is most common, should remain vigilant regarding potential complications during pregnancy, given the implications for both maternal and neonatal health.

The aim of the present study was to synthesize and describe pregnancy outcomes and neonatal birth weights were reported in case reports over the past 30 years. By summarizing the clinical features, complications, and delivery characteristics of pregnancies complicated by BD, this review seeks to provide obstetricians with practical insights that may assist in managing pregnancy and delivery in BD and related autoimmune vasculitides.

The hypothesis of this study is that pregnancies complicated by BD are associated with a higher frequency of adverse maternal and neonatal outcomes, particularly preterm delivery and low birth weight.

METHODS

This study is based on a review of published case reports and does not require an ethics committee report because it does not involve direct patient recruitment. All procedures were carried out in accordance with the ethical rules and the principles.

This study was designed as a descriptive pooled analysis of published case reports and small case series on pregnancy outcomes in BD. A structured and transparent literature search was performed, adapted for case-based data. Two independent reviewers screened eligible studies, assessed report completeness, and extracted maternal and neonatal

Corresponding Author: Erhan Muluk, my48523@gmail.com



This work is licensed under a Creative Commons Attribution 4.0 International License.

data for pooled descriptive analysis. No meta-analysis was conducted.

A comprehensive literature search was carried out and adapted for case-based data. PubMed, the Web of Science Core Collection, and Google Scholar were searched for reports published between January 1995 and March 2025.

In PubMed, the following query was applied: [Behcet (Title)] AND [pregnancy (Title) OR pregnant (Title)], with results subsequently filtered to include only case reports. In the Web of Science Core Collection, two separate searches were performed using the following strings: pregnancy AND Behcet AND "case report" (topic) and pregnant AND Behcet AND "case report" (topic). In addition, "case series" was used in place of "case report" to capture small series relevant to the topic. A complementary search using the same terms was also carried out in Google Scholar. Reference lists of the retrieved papers were also reviewed to identify any additional eligible cases.

The author and another independent reviewer screened all records, removed duplicates, and assessed eligibility using predefined criteria. Studies were included if they described pregnancy in a woman with a confirmed diagnosis of BD, reported maternal and/or neonatal outcomes, were published in English, and appeared within the specified time frame. Reports were excluded if they lacked outcome data, were reviews, conference abstracts, or experimental studies, described non-pregnant or male patients, or duplicated previously published cases. Any differences in judgment between reviewers were resolved by discussion and agreement.

Data from each eligible report were extracted independently by both reviewers using a standardized form. Recorded variables included maternal age, disease duration, disease activity during pregnancy, mode of delivery, and neonatal birth weight. When available, additional information such as miscarriage, preeclampsia, or preterm delivery was also noted. The completeness of each report was assessed, and studies that did not provide sufficient clinical or outcome details were excluded from the pooled analysis.

Statistical Analysis

Descriptive statistics were used to summarize the findings. Categorical data were expressed as numbers and percentages, and continuous data as mean±standard deviation (SD). All analyses were performed using Microsoft Excel 2021.

RESULTS

The systematic search retrieved a total of 73 articles. After removing duplicates and excluding publications that were irrelevant, lacked outcome data, or did not provide clear case descriptions, 61 cases with sufficient clinical detail were included in the pooled analysis (Figure 1).⁸⁻²⁵ Most of the single cases were published in year 2025 (Figure 2). Data availability varied across variables. Maternal age was reported in 58/61 cases, disease duration in 52/61, disease activity in 61/61, and neonatal birth weight in 54/61 cases. Missing data were mainly due to incomplete clinical descriptions in older reports.

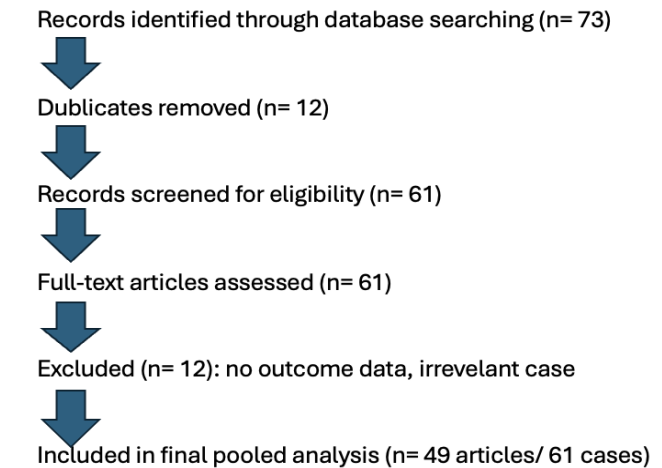


Figure 1. Flow of study selection and inclusion process

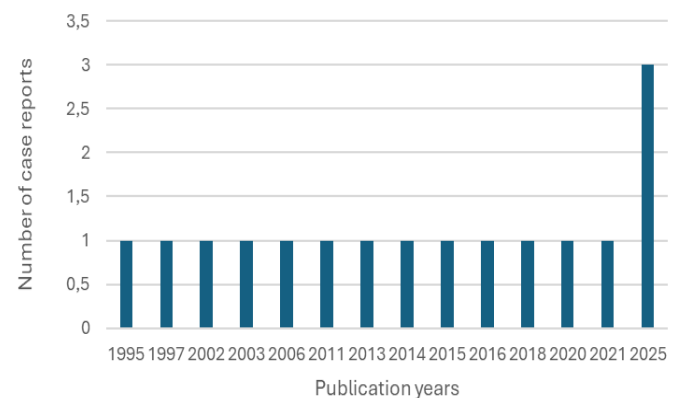


Figure 2. Distribution of single case reports of Behçet's disease in pregnancy by year

The mean maternal age at pregnancy was 32.33 ± 6.34 years (range 19-45), and the mean disease duration was 5.25 ± 3.22 years (range 0-11). Disease activity during pregnancy was reported in 20 of 61 cases, corresponding to a flare rate of 32.8%.

Adverse pregnancy outcomes were observed in a subset of patients. Preterm delivery occurred in 7 cases (11.5%), while miscarriage and preeclampsia were reported in 2 (3.3%) and 1 (1.6%) cases, respectively. Low birth weight (<2500 g) was documented in 7 infants (11.5%). The mean neonatal birth weight across all reported cases was 3045 ± 517 g.

A summary of extracted variables is presented in Table, and the distribution of complications is illustrated in Figure 3.

Table. Maternal and neonatal outcomes extracted from published case reports and case series of pregnancy in Behçet's disease (1995-2025)

Variable	Mean (SD)/n (percentage)
Maternal age	32.33 (SD: 6.34)
Disease duration	5.25 (SD: 3.22)
Flare at pregnancy	20/61 (32.79%)
Preterm delivery	7/61 (11.48%)
Miscarriage	2/61 (3.28%)
Preeclampsia	1/61 (1.64%)
Low birth weight	7/61 (11.48%)
Neonatal birth weight (g)	3045 (SD: 517)

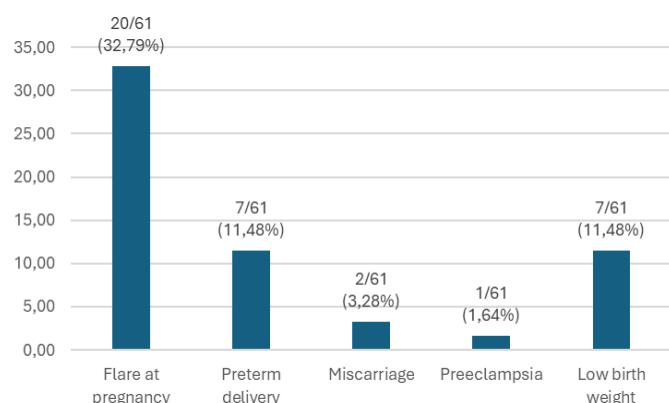


Figure 3. Distribution of pregnancy complications among women with Behçet's disease reported in published case reports (1995-2025)

Of the 61 included cases, the majority were reported from Türkiye (n=37), followed by Japan (n=15). Other cases originated from the United States (n=3), Iran (n=1), Italy (n=1), Portugal (n=1), France (n=1), Bulgaria (n=1), and China (n=1). This distribution reflects the geographic clustering of BD along the Silk Road (Figure 4). The cases reported from the United States are likely to have occurred predominantly among immigrants from high-prevalence regions.

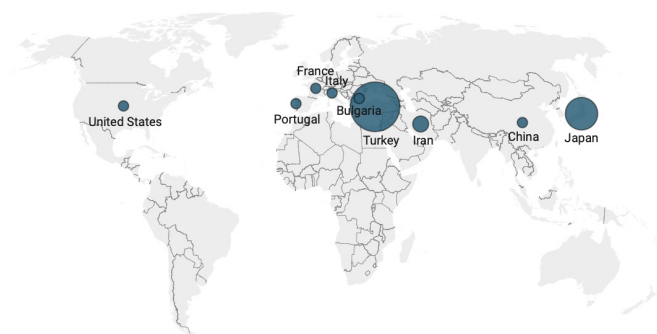


Figure 4. Geographic distribution of Behçet disease cases (1995-2025)

DISCUSSION

This study pooled published case reports of pregnancies complicated by BD over the past three decades. The analysis of 61 unique cases showed that approximately one-third of women experienced disease flares during pregnancy, while the most frequent adverse obstetric outcomes were preterm delivery (11.5%) and low birth weight (11.5%). Miscarriage and preeclampsia were reported less frequently. These findings, although derived from case reports with inherent limitations, are in line with observations from larger cohort studies, suggesting that BD may predispose women to certain pregnancy complications.

Of note, some of the included women had a disease duration of zero years, meaning that BD was first diagnosed during pregnancy. This highlights an important clinical point: obstetricians should remain alert to the possibility of BD first presenting in pregnancy, particularly when confronted with unexplained mucocutaneous lesions, ocular symptoms, or systemic inflammatory features.

When comparing our pooled results with data from larger cohorts, the pattern is consistent. In a nationwide retrospective

cohort study from South Korea including more than 12 million births, the prevalence of BD in pregnancy was 1.14 per 100,000 births. Among 144 women with BD identified in that cohort, hospitalization for non-obstetric reasons was nearly twice as common, and affected women had a higher risk of preterm birth and postpartum venous thromboembolism. Their infants were also more likely to be born prematurely.²⁶ These data reinforce the notion that preterm delivery and low birth weight are recurrent themes in BD pregnancies, even beyond case reports.

In our analysis, exacerbations during pregnancy occurred in 20/61 cases (32.8%). Although it is difficult to establish a causal relationship, active disease may increase the risk of preterm birth and low birth weight through inflammatory and thrombotic processes affecting the uteroplacental circulation. This interpretation is consistent with reports showing that vasculitides increase pregnancy loss and preterm birth among affected women compared with the general obstetric population.²⁷ In BD specifically, placental involvement driven by inflammatory and thrombotic processes has been suggested as a major contributor to adverse outcomes.²⁸ Likewise, an increased risk of preterm birth and postpartum venous thromboembolism in pregnant women with BD has been reported.²⁶ We did not find a direct association between exacerbations and adverse outcomes in our dataset, primarily due to the small sample size, heterogeneity among cases, and underreporting of some variables. Therefore, larger, prospective studies are needed to better evaluate the relationship between disease activity and pregnancy outcomes.

Another important consideration is medical management. Many of the drugs used in BD including corticosteroids, azathioprine, cyclosporine, colchicine, and certain biologic agents have varying levels of safety data in pregnancy. Corticosteroids remain the mainstay for controlling flares and are generally considered safe when used at the lowest effective dose. Azathioprine and cyclosporine have also been used successfully when disease activity requires maintenance therapy, while colchicine appears relatively safe based on data from familial Mediterranean fever. In contrast, methotrexate and mycophenolate mofetil are contraindicated because of their teratogenic potential and must be discontinued before conception. Experience with anti-TNF agents such as infliximab and adalimumab is growing, and several reports describe successful pregnancies with their continued use into the second trimester. Drug selection should be individualized and ideally guided by a rheumatologist familiar with current recommendations. When disease activity is severe, the risks of untreated inflammation often outweigh potential medication risks. After delivery, breastfeeding can usually be encouraged, as most of these drugs appear in breast milk only in minimal amounts.²⁷⁻³⁰

General principles established for more common autoimmune rheumatic diseases are also applicable to BD: pregnancies should ideally be planned during periods of disease remission; patients must be counseled thoroughly before conception; regular monitoring during gestation and the postpartum period is crucial; and multidisciplinary management involving rheumatologists, obstetricians, and neonatologists

is recommended. In the event of relapse, timely and sometimes intensive treatment should not be delayed, as poorly controlled disease poses greater risks than most therapies.²⁹

In summary, although based on pooled case reports, our findings underscore that preterm delivery and low birth weight are the most frequent complications of BD pregnancies. Obstetricians should be mindful that BD may first manifest during pregnancy, and ongoing collaboration with rheumatologists is essential for safe maternal and neonatal outcomes.

Another point to consider is the geographic distribution of the reported cases. In our dataset, most cases originated from Türkiye and Japan, reflecting the known clustering of BD along the historic Silk Road. This uneven distribution likely mirrors regional disease prevalence but may also introduce selection and publication bias. Physicians in high-prevalence regions may be more experienced in recognizing and reporting BD during pregnancy, whereas cases from low-prevalence areas might be underrepresented or published less frequently. Therefore, the predominance of cases from certain countries could have influenced the apparent frequency of pregnancy complications and limits the generalizability of these findings.

Limitations

This study has several limitations that must be acknowledged. First, the analysis was based exclusively on published case reports and case series, which are inherently subject to reporting bias and often emphasize unusual or complicated cases. As a result, the true frequency of pregnancy complications in BD cannot be reliably estimated from these data alone. Second, clinical information was not uniformly available across all reports, and missing data may have influenced the pooled results. Third, management strategies, diagnostic criteria, and treatment availability have changed over the three decades covered by this study, introducing heterogeneity that could not be fully controlled. Finally, because no control group was included, direct comparison with the general obstetric population was not possible within this dataset.

CONCLUSION

In this study, preterm delivery and low birth weight emerged as the most frequent pregnancy complications among women with BD. Obstetricians should provide closer monitoring of affected women, with particular attention to signs of preterm labor and fetal growth restriction. Counseling patients before conception, ensuring multidisciplinary management with rheumatologists, and planning pregnancies during remission are essential steps to optimize maternal and neonatal outcomes. Physicians practicing in high-prevalence regions such as Türkiye and Japan should be especially aware of the disease burden. Larger, well-designed cohort studies are needed to better define risks, guide treatment strategies, and improve outcomes for both mothers and their infants.

ETHICAL DECLARATIONS

Ethics Committee Approval

This study is based on a review of published case reports and does not require an ethics committee report because it does not involve direct patient recruitment.

Informed Consent

Not applicable, since all data were obtained from previously published reports.

Referee Evaluation Process

Externally peer-reviewed

Conflict of Interest Statement

The author has no conflicts of interest to declare.

Financial Disclosure

The author declared that this study has received no financial support.

Author Contributions

Conceptualization, methodology, and literature search; data curation and formal analysis; visualization, writing-original draft preparation, and writing-review and editing: Erhan Muluk.

Acknowledgements

The author thanks Prof Dr. Özer Öztekin for being the second independent reviewer during the literature screening and data extraction stages of this study.

REFERENCES

- Hatemi G, Uçar D, Uygunoğlu U, Yazici H, Yazici Y. Behçet syndrome. *Rheum Dis Clin North Am*. 2023;49(3):585-602. doi:10.1016/j.rdc.2023.03.010
- Emmi G, Bettiol A, Hatemi G, Prisco D. Behçet's syndrome. *Lancet*. 2024;403(10431):1093-1108. doi:10.1016/S0140-6736(23)02629-6
- Merlino L, Del Prete F, Loboizzo B, Priori R, Piccioni MG. Behçet's disease and pregnancy: a systematic review. *Minerva Ginecol*. 2020;72(5):332-338. doi:10.23736/S0026-4784.20.04564-5
- Jadaon J, Shushan A, Ezra Y, Sela HY, Ozcan C, Rojansky N. Behçet's disease and pregnancy. *Acta Obstet Gynecol Scand*. 2005;84(10):939-944. doi:10.1111/j.0001-6349.2005.00761.x
- Ben-Chetrit E. Behçet's syndrome and pregnancy: course of the disease and pregnancy outcome. *Clin Exp Rheumatol*. 2014;32(4 Suppl 84):S93-98.
- Davatchi F, Chams-Davatchi C, Shams H, et al. Behçet's disease: epidemiology, clinical manifestations, and diagnosis. *Expert Rev Clin Immunol*. 2017;13(1):57-65. doi:10.1080/1744666X.2016.1205486
- Maldini C, Druce K, Basu N, LaValley MP, Mahr A. Exploring the variability in Behçet's disease prevalence: a meta-analytical approach. *Rheumatology (Oxford)*. 2018;57(1):185-195. doi:10.1093/rheumatology/kew486
- Takayama K, Ishikawa S, Enoki T, Kojima T, Takeuchi M. Successful treatment with infliximab for Behçet disease during pregnancy. *Ocul Immunol Inflamm*. 2013;21(4):321-323. doi:10.3109/09273948.2013.781655
- Fujikawa K, Endo Y, Mizokami A, et al. Successful Treatment with Adalimumab for Intestinal Behçet's disease during Pregnancy. *Intern Med*. 2016;55(10):1375-1378. doi:10.2169/internalmedicine.55.6590
- Mainini G, Di Donna MC, Esposito E, et al. Pregnancy management in Behçet's disease treated with uninterrupted infliximab. Report of a case with fetal growth restriction and mini-review of the literature. *Clin Exp Obstet Gynecol*. 2014;41(2):205-207.
- Li X, Huang X, Yu X. Behçet's disease in pregnancy: a case report and literature review. *BMC Pregnancy Childbirth*. 2025;25(1):69. doi:10.1186/s12884-024-07063-0
- Gomes V, Mesquita A, Capela C. Autoimmune diseases and pregnancy: analysis of a series of cases. *BMC Res Notes*. 2015;8:216. doi:10.1186/s13104-015-1177-x
- Buntrock A, McNamara M, Tantoush H, Von Wald T. Worsening of Behçet syndrome symptoms during pregnancy. *S D Med*. 2021;74(3):112-114.

14. Fotaki P, Deligeoroglou E, Michailidis E, et al. Adamantiades-Behcet's syndrome and pregnancy: a case report. *Akush Ginekol (Sofia)*. 2002; 42(1):38-39.
15. Mirfeizi Z, Memar B, PourZand H, Molseghi MH, Rezaei Shahmirzadi A, Abdolahi N. Ventricular endomyocardial fibrosis in a pregnant female with Behçet's disease. *Asian Cardiovasc Thorac Ann*. 2018;26(8):619-621. doi:10.1177/0218492316687177
16. Nishii S, Yamaguchi K, Amita M, Saito T, Saito H, Sekizawa A. Successful pregnancy in a case of Behçet's disease after treatment with prednisolone. *Case Rep Obstet Gynecol*. 2020;2020:8862651. doi:10.1155/2020/8862651
17. Guzelian G, Norton ME. Behçet's syndrome associated with intrauterine growth restriction: a case report and review of the literature. *J Perinatol*. 1997;17(4):318-320.
18. Xie L, Benson T, Morgan EE, et al. New onset Behçet's disease in repeat pregnancy. *JAAD Case Rep*. 2025;58:106-108. doi:10.1016/j.jdc.2024.10.040
19. Kale A, Akyildiz L, Akdeniz N, Kale E. Pregnancy complicated by superior vena cava thrombosis and pulmonary embolism in a patient with Behcet disease and the use of heparin for treatment. *Saudi Med J*. 2006;27(1):95-97.
20. Wechsler B, Génèreau T, Biousse V, et al. Pregnancy complicated by cerebral venous thrombosis in Behçet's disease. *Am J Obstet Gynecol*. 1995;173(5):1627-1629. doi:10.1016/0002-9378(95)90663-0
21. Cakal B, Koklu S, Beyazit Y, Ozdemir A, Beyazit F, Ulker A. Fatal colonic perforation in a pregnant with Behçet's disease. *J Crohns Colitis*. 2011;5(3):273-274. doi:10.1016/j.crohns.2011.03.003
22. Duran B, Şalk S, Erden O, Çetin M, Demirköprülü N. Vulvar oedema associated with Behçet's disease in pregnancy: a case report. *Türkiye Klinikleri J Gynecol Obst*. 2003;13(6):463-465.
23. Shimada H, Wakiya R, Nakashima S, et al. Clinical courses and pregnancy outcomes of twelve cases complicated with Behcet's disease in our institution. *Ann Rheum Dis*. 2023;82(Suppl 1):1974. doi:10.1136/annrheumdis-2023-eular.5190
24. Demir O, Ömeroğlu Demir G. Behçet's disease diagnosed with a multidisciplinary approach in pregnancy. *Bezmialem Science*. 2025; 13(1):81-83. doi:10.14235/bas.galenos.2024.81904
25. Erenel H, Alici Davutoğlu E, Özel A, Karslı F, Korkmaz SÖ, Madazlı R. Behçet's disease and pregnancy: obstetric outcomes of 33 patients. *Sisli Etfal Hastan Tip Bul*. 2017;51(4):318-321. doi:10.5350/SEMB.20170725042231
26. Lee S, Czuzoj-Shulman N, Abenhaim HA. Behcet's disease and pregnancy: obstetrical and neonatal outcomes in a population-based cohort of 12 million births. *J Perinat Med*. 2019;47(4):381-387. doi:10.1515/jpm-2018-0161
27. Machen L, Clowse ME. Vasculitis and pregnancy. *Rheum Dis Clin North Am*. 2017;43(2):239-247. doi:10.1016/j.rdc.2016.12.005
28. Orgul G, Aktoz F, Beksac MS. Behcet's disease and pregnancy: what to expect? *J Obstet Gynaecol*. 2018;38(2):185-188. doi:10.1080/01443615.2017.1336614
29. Doria A, Iaccarino L, Ghirardello A, et al. Pregnancy in rare autoimmune rheumatic diseases: UCTD, MCTD, myositis, systemic vasculitis and Behcet's disease. *Lupus*. 2004;13(9):690-695. doi:10.1191/0961203304lu1098oa
30. Pagnoux C. Pregnancy and vasculitides. *Presse Med*. 2008;37(11):1657-1665. doi:10.1016/j.lpm.2008.06.016