ACTA MEDICA
formerly Hacettepe Medical Journal

www.actamedica.org

Volume 57
Issue 3
September 2026



Online ISSN: 2147-9488

from the seniors to the students



ACTA MEDICA

formerly Hacettepe Medical Journal

Volume **57** • Issue **3** • September **2026**

Online ISSN: 2147-9488

www.actamedica.org

ACTA MEDICA
formerly Hacettepe Medical Journal

Acta Medica

ISSN (Online): 2147-9488

www.actamedica.org

Cilt 57, Sayı 3, Eylül 2026

Sahibi
Hacettepe Üniversitesi
Tıp Fakültesi adına sahibi
Hakan Göker

Sorumlu Yazı İşleri Müdürü
Gözde Yazıcı

Yayının Türü
Yaygın Süreli Yayının

Yayın Sıklığı ve Dili
Üç aylık, İngilizce

Baş Editörler
Gözde Yazıcı
Sevinç Sarınç

Editöryal İletişim
Hacettepe Üniversitesi
Tıp Fakültesi Dekanlığı
06100 Sıhhiye - Ankara, Türkiye
E-posta: editor@actamedica.org

Yayıncı
Hacettepe Üniversitesi
Tıp Fakültesi Dekanlığı
06100 Sıhhiye - Ankara, Türkiye
Telefon: 0 312 305 10 80
Belgeç (faks): 0 312 310 05 80
E-posta: tipmaster@hacettepe.edu.tr

Yayıncılık Hizmetleri
Akdema Bilişim ve Yayıncılık
E-posta: bilgi@akdema.com
Web: www.akdema.com

Acta Medica

ISSN (Online): 2147-9488

www.actamedica.org

Volume 57, Issue 3, September 2026

Owner
Owner on behalf of the Hacettepe University
Faculty of Medicine
Hakan Göker

Administrator
Gözde Yazıcı

Publication Type
Peer-reviewed journal

Publication Frequency and Language
Quarterly, English

Editors-in-Chief
Gözde Yazıcı
Sevinç Sarınç

Editorial Office
Hacettepe University
Faculty of Medicine
06100 Sıhhiye - Ankara, Türkiye
Email: editor@actamedica.org

Publisher
Hacettepe University
Faculty of Medicine
06100 Sıhhiye - Ankara, Türkiye
Phone: 0 312 305 10 80
Fax: 0 312 310 05 80
Email: tipmaster@hacettepe.edu.tr

Publishing Services
Akdema Informatics and Publishing
Email: bilgi@akdema.com
Web: www.akdema.com

Acta Medica is an peer-reviewed, open access journal published by the Hacettepe University Faculty of Medicine. All articles are published under the terms of the **Creative Commons Attribution License (CC-BY)** which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

You can reach all publication policies and author guidelines from **actamedica.org**

ACTA MEDICA
formerly Hacettepe Medical Journal

Editorial Board

Administrator

- **Gözde Yazıcı**, MD, Department of Medical Sciences, Hacettepe University, Ankara, Türkiye

Editors-in-Chief

- **Gözde Yazıcı**, MD, Department of Medical Sciences, Hacettepe University, Ankara, Türkiye
- **Sevinç Sarınç**, MD, Department of Medical Sciences, Hacettepe University, Ankara, Türkiye

Editors

- **Burak Yasin Aktaş**, MD, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Yavuz Ayhan**, MD, Department of Psychiatry, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Derman Başaran**, MD, Department of Obstetrics and Gynecology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Nursel Çalık Başaran**, MD, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Pınar Çalış**, MD, Department of Obstetrics and Gynecology, Faculty of Medicine, Gazi University, Ankara, Türkiye
- **Başak Çeltikçi**, MD, PhD, Department of Medical Biochemistry, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Hemra Çil**, MD, Department of Anesthesia and Perioperative Care, University of California, California, USA
- **Saniye Ekinci**, MD, Department of Pediatric Surgery, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Güneş Esendağlı**, PhD, Basic Oncology Subdivision, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Volkan Genç**, MD, Department of General Surgery, Faculty of Medicine, Ankara University, Ankara, Türkiye
- **Güneş Güner**, MD, Department of Medical Pathology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Ekim Gümeler**, MD, Department of Radiology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Ahmet Çağkan İnkaya**, MD, Department of Infectious Diseases and Clinical Microbiology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Murat İzgi**, MD, PhD, Department of Anesthesiology and Reanimation, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Emre Kara**, Phar, PhD, Department of Clinical Pharmacy, Faculty of Pharmacy, Hacettepe University, Ankara, Türkiye
- **Murat Kara**, MD, Department of Plastic Reconstructive and Aesthetic Surgery, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Zeynep Ceren Karahan**, MD, Department of Medical Microbiology, Faculty of Medicine, Ankara University, Ankara, Türkiye
- **Saygın Kamacı**, MD, Department of Orthopedics and Traumatology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **İrem Koç**, MD, Department of Ophthalmology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Orhan Murat Koçak**, MD, Department of Mental Health and Diseases, Faculty of Medicine, Başkent University, Ankara, Türkiye
- **Dilek Menemenlioğlu**, MD, Vaccine Institute, Hacettepe University, Ankara, Türkiye

- **Müge Yemişci Özkan**, MD, PhD, Institute of Neurological Sciences and Psychiatry, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Ahmet Erim Pamuk**, MD, Department of Otorhinolaryngology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Esra Serdaroğlu**, MD, Department of Pediatrics, Faculty of Medicine, Gazi University, Ankara, Türkiye
- **Süleyman Nahit Şendur**, MD, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Yeşim Er Öztaş**, MD, Department of Medical Biochemistry, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Murat Sincan**, MD, Department of Internal Medicine, The University of South Dakota, USA
- **Gülşen Taşdelen Teker**, PhD, Department of Medical Education and Informatics, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **İdil Rana User**, MD, Department of Pediatric Surgery, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Oğuz Abdullah Uyaroğlu**, MD, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Şule Ünal**, MD, Department of Pediatrics, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
- **Tolga Yıldırım**, MD, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Language Editor

- **Sinem Akgül**, MD, Department of Medical Sciences, Hacettepe University, Ankara, Türkiye
- **Başak Çeltikçi**, MD, PhD, Department of Basic Sciences, Hacettepe University, Ankara, Türkiye

Statistics Editor

- **Sevilay Karahan**, PhD, Department of Basic Sciences, Hacettepe University, Ankara, Türkiye

ACTA MEDICA
formerly Hacettepe Medical Journal

Contents

Research Articles

Evaluation of the relation between serum 25-OH vitamin D levels and clinical findings of febrile neutropenia in childhood cancer patients
Burçe Orman, Gürses Şahin 187

Evaluating the proficiency of ChatGPT-3.5, ChatGPT-4o and Google Gemini in responding to frequently asked questions on gender affirming care by transgender people
Selin Tekin, Banu Ertürk, Seda Hanife Oğuz, Bülent Okan Yıldız 193

Clinical characteristics and legal outcomes of orthopedic malpractice claims in the constitutional court of Türkiye
Barış Kafa, Ulaş Can Kolaç, Mazlum Veysel Sili 202

Quantitative CT assessment of thymic involution in adults: Role of age, sex and smoking
Selin Ardalı Düzgün, İlke Taşçı, Jale Karakaya, Gamze Durhan 211

Recognizable clinical features of 1p36 deletion syndrome: Insights from a single-center cohort
Gizem Ürel Demir, Serap Ketenci İşlek, Gülen Eda Utine, Pelin Özlem Şimşek Kiper 220

Determinants of functional recovery after tibial plateau fractures: The role of fracture severity, fixation, and complexity
Gökhan Ayık, Berat Yüksel, Orhan Mete Karademir, Ulaş Can Kolaç 227

Comparison of serum oxidative stress biomarkers in amateur football players and sedentary men
Müslüm Gök, Kazım Can Aytamur, Mustafa Utku Sarı, Serkan Aksu, Gönül Babayiğit İrez, Kemal Göral 238

Distinct psychological correlates of burnout dimensions among healthcare workers
Talat Demirsöz, Mevhibe İrem Yıldız 246

Comparative genomic profiling of early-onset and late-onset colorectal cancer
Bengi Öztürk, Hatime Arzu Yaşar, Onur Keskin, Taylan Kav 258

Outcomes of oral provocation tests with alternative iron preparations in patients with suspected iron hypersensitivity
Özge Öztürk Aktaş, Ebru Damadoğlu, Gül Karakaya, Ali Fuat Kalyoncu 266

Intensive care unit admissions in adult cystic fibrosis patients: A national cohort study from Türkiye
Mehmet Yıldırım, Emin Gemcioğlu, Rıfat Bozkuş, Şuayip Birinci, Arzu Topeli, Naim Ata 273

Improvement in geriatric syndromes following clinical response to endoscopic third ventriculostomy in idiopathic normal pressure hydrocephalus
Yasemin Polat, Ömer Özpişkin, Merve Güner, Eyüp Bayatlı, Onur Özgüral, Gökmen Kahiloğulları, Mustafa Ağahan Ünlü, Mert Eşme, Cafer Balcı, Meltem Gülhan Halil, Mustafa Cankurtaran, Burcu Balam Doğu 283

Letters to the Editor

When can antiseizure medications be safely withdrawn after encephalitis?
Okan Sökmen 293

Letter to editor: Presentation and clinical outcome of adolescents with graves' disease
İrem Şenoymak, Mustafa Can Şenoymak 296

Acta Medica 2026; 57(3): 187-192 DOI: 10.32552/actamedica.2026.1120

acta medica RESEARCH ARTICLE

Evaluation of the relation between serum 25-OH vitamin D levels and clinical findings of febrile neutropenia in childhood cancer patients

Burçe Orman[1], Gürses Şahin[2]

[1]Department of Pediatric Endocrinology, University of Health Sciences, Dr. Sami Ulus Research and Education Hospital, Ankara, Türkiye
[2]Department of Pediatric Oncology, University of Health Sciences, Dr. Sami Ulus Research and Education Hospital, Ankara, Türkiye

Abstract

Objective: This retrospective study aimed to evaluate the association between serum 25-OH vitamin D (25OHD) levels and the clinical course of febrile neutropenia in children with cancer.

Materials and Methods: Seventy-seven pediatric cancer patients with febrile neutropenia and 25OHD levels measured at diagnosis were included. Forty-four healthy children served as controls. Clinical data including febrile neutropenia duration and related parameters were retrospectively reviewed.

Results: The study included 77 patients (44 male, 33 female) and 44 controls (29 male, 15 female). Spearman correlation analysis revealed a significant negative correlation between 25OHD levels and neutropenia duration (r=-0.249, p=0.029) and number of febrile days (r=-0.253, p=0.027). However, multivariate regression analysis showed that 25OHD level was not an independent predictor of these outcomes; leukemia diagnosis was the only independent predictor (p<0.001).

Conclusion: Lower 25OHD levels were associated with longer neutropenia duration and more febrile days. However, this association was not independent of diagnosis. Vitamin D monitoring in pediatric cancer patients may be clinically relevant, but prospective studies are needed to confirm its role.

Keywords: children, cancer, febrile neutropenia, vitamin D, immunomodulation

Introduction

Vitamin D regulates cell differentiation and proliferation through its nuclear receptor. Beyond its role in bone metabolism, it has anti-inflammatory and immunomodulatory properties with effects on cancer, autoimmunity, infections, and cardiovascular health [1-3].

Deficiency is linked to increased risk of viral respiratory infections and sepsis [4]. Vitamin D receptors are expressed on T cells, B cells, macrophages, and dendritic cells, suggesting a direct role in innate and adaptive immunity.

Febrile neutropenia causes 5% mortality in solid tumors and up to 10% in hematological malignancies. It is a critical, treatment-limiting complication in pediatric oncology [5]. Pediatric cancer patients are at high risk for vitamin D deficiency due to reduced sun exposure, chemotherapy-related malabsorption, and increased metabolic demands.

Corresponding author: Burçe Orman • **Email:** burceayvazoglu@gmail.com

Received: January 14, 2025 **Accepted:** August 16, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Vitamin D deficiency is prevalent in pediatric oncology patients and may negatively affect clinical outcomes [6,7]. However, its specific impact on febrile neutropenia in children with cancer is not well established. This study aimed to investigate this association.

Materials and Methods

Study design and patients

This was a retrospective, observational, single-center study. The patient group consisted of children aged 0-18 years hospitalized in the pediatric oncology clinic of Dr. Sami Ulus Research and Education Hospital between January 2014 and January 2018, with 25OHD levels measured at cancer diagnosis.

The control group included 44 healthy children without chronic disease and was used to compare baseline 25OHD levels with those of cancer patients at diagnosis, providing a reference for the prevalence of vitamin D deficiency in this population.

The study focused on patients who did not receive vitamin D supplementation during the first two months after diagnosis but experienced febrile neutropenia. We analyzed the timing of febrile neutropenia onset, its severity, and its duration.

Definitions

Febrile conditions were defined as a body temperature >38°C measured via the tympanic route. Neutropenia was defined as an absolute neutrophil count (ANC) <500/mm^3, or an ANC <1500/mm^3 predicted to fall further. Neutropenia severity was classified as: severe (ANC <500/mm^3), moderate (ANC 500-1000/mm^3), or mild (ANC 1000-1500/mm^3).

25OHD levels were measured using liquid chromatography. Vitamin D status was classified as: deficient (≤15 ng/ml), insufficient (15-20 ng/ml), sufficient (20-100 ng/ml), and toxic (>150 ng/ml) [8]. The cut-off of ≤15 ng/ml for deficiency was chosen based on the classification recommended by the American Academy of Pediatrics and used in Turkish pediatric guidelines at the time of the study. Although a threshold of 20 ng/ml is more commonly used in recent literature, our cut-off reflects the standard applied in clinical practice during the study period. None of the patients were receiving vitamin D supplements at the time of febrile neutropenia episode.

Data collection

The duration of neutropenia, number of days with fever, presence of complications (e.g., mucositis, abscesses), use of glycopeptides and antifungals, and G-CSF (granulocyte colony-stimulating factor) use were recorded from patient files.

Statistical analysis

Data were analyzed using SPSS (Statistical Package for the Social Sciences). Normality was assessed using the Shapiro-Wilk test. As all continuous variables showed non-normal distribution ($p<0.05$), non-parametric tests were used throughout. The Mann-Whitney U test was used for comparisons between two groups, and the Kruskal-Wallis test for comparisons among three or more groups. Spearman correlation analysis was performed to assess the relationship between 25OHD levels and continuous outcome variables. Multivariate linear regression analysis was conducted to identify independent predictors of neutropenia duration and number of febrile days, including 25OHD level, diagnosis (leukemia vs. other), sex, and G-CSF use as covariates. Results are reported with beta coefficients, standard errors, 95% confidence intervals, and p-values. Fisher's Exact test was used for categorical data. Statistical significance was set at $p<0.05$.

Ethical approval

The study was approved by the Health Sciences University Ankara Pediatrics Hematology and Oncology Training and Research Hospital Ethics Committee (approval number: 2018-100, date: 09.07.2018) and was conducted in accordance with the Declaration of Helsinki. This study was conducted retrospectively through review of patient medical records.

Results

There were 77 children (44 male, 33 female) in the patient group and 44 children (29 male, 15 female) in the control group. The mean age of the patient group was 7.99±5.1 years, while the control group had a mean

age of 8.36±2.48 years (p=0.197). Among the patients, 42 had leukemia (54.5%), 13 had lymphoma (16.9%), and 22 had solid tumors (28.6%).

The mean 25OHD level in the patient group was 15.82±8.63 ng/ml, while in the control group it was 18.08±9.41 ng/ml (p=0.060). Vitamin D deficiency (25OHD <20 ng/ml) was found in 63.6% (n=49) of patients and 47.7% (n=21) of controls. Among the patients, 66.6% (n=28/42) with leukemia and 60% (n=21/35) with lymphoma or solid tumors had insufficient 25OHD levels (p=0.120).

The mean white blood cell count was 6182±1091 $\times 10^3$/µL, the mean ANC at onset of febrile neutropenia was 938±468 $\times 10^3$/µL, and the mean ANC at recovery was 2432±916 $\times 10^3$/µL. The mean duration of neutropenia was 19.83±16.65 days, and the mean number of days with fever was 4.66±2.62 days.

In patients who received vitamin D prophylaxis during their first year of life (18.2%), the mean 25OHD level was 27.79±7.39 ng/ml, compared to 13.16±6.36 ng/ml in those who did not (81.8%) (p<0.001).

As 25OHD levels decreased, both the duration of neutropenia and the number of days with fever increased significantly (p<0.05) (Table 1). Patients with severe neutropenia (ANC <500/mm^3) had the lowest mean 25OHD levels (13.45±7.76 ng/ml) compared to moderate (15.88±10.88 ng/ml) and mild groups (16.69±7.81 ng/ml), although this did not reach statistical significance (p=0.164).

In patients with leukemia, neutropenia duration was significantly longer (26.01±19.65 vs 12.43±7.15 days; p=0.031) and febrile days were more frequent (5.71±2.91 vs 3.39±2.25 days; p=0.032) compared to non-leukemic patients (Table 2).

Table 1. Clinical evaluation based on 25-OH vitamin D levels.

	<15 ng/ml mean±SD	15-20 ng/ml mean±SD	>20 ng/ml mean±SD	p-value
Duration of neutropenia (days)	22.40±16.44	18.00±15.49	16.43±17.15	0.044*
Fever (°C)	38.4±0.43	38.6±0.48	38.6±0.41	0.059
Number of febrile days	5.14±2.27	5.12±3.13	3.86±2.89	0.015*
WBC (10^3/µL)	5859.52±9747.02	3402.86±1625.92	7363.21±13679.91	0.744
ANC at onset (10^3/µL)	903.10±468.27	1227.14±234.36	919.64±415.09	0.205
ANC at recovery (10^3/µL)	2289.52±643.39	2742.86±945.60	2570.71±1208.70	0.392

ANC: Absolute neutrophil count, WBC: White blood cell. *Statistically significant (p<0.05). p-values were calculated using the Kruskal-Wallis test.

Table 2. Clinical evaluation according to diagnosis.

	Leukemia (n=42 mean±SD	Others (n=35) mean±SD	p-value
25OHD level (ng/ml)	15.18±9.00	16.59±8.24	0.481
Duration of neutropenia (days)	26.00±19.65	12.43±7.15	0.031*
Fever (°C)	38.57±0.44	38.49±0.44	0.473
Number of febrile days	5.71±2.91	3.39±2.25	0.032*
WBC (10^3/µL)	7833.57±14420.57	4202.29±2887.73	0.118
ANC at onset (10^3/µL)	985.95±452.33	881.71±420.84	0.302
ANC at recovery (10^3/µL)	2406.90±761.93	2464.29±1084.46	0.786

ANC: Absolute neutrophil count, WBC: White blood cell, 25OHD: 25-hydroxy vitamin D. *Statistically significant (p<0.05). p-values were calculated using the Mann-Whitney U test.

Table 3. Multivariate linear regression analysis for predictors of neutropenia duration and number of febrile days.

Variable	Beta	SE	p-value	95% CI
Neutropenia duration (days)				
25OHD level	-0.289	0.209	0.171	-0.706 to 0.127
Leukemia diagnosis	14.538	3.479	<0.001*	7.603 to 21.474
Sex (male)	2.442	3.615	0.502	-4.765 to 9.650
G-CSF use	-3.803	3.719	0.310	-11.217 to 3.612
Number of febrile days				
25OHD level	-0.047	0.033	0.160	-0.112 to 0.019
Leukemia diagnosis	2.341	0.546	<0.001*	1.251 to 3.430
Sex (male)	0.234	0.568	0.681	-0.898 to 1.366
G-CSF use	-0.419	0.584	0.475	-1.584 to 0.745

Spearman correlation analysis demonstrated a significant negative correlation between 25OHD levels and neutropenia duration (r=-0.249, p=0.029) and number of febrile days (r=-0.253, p=0.027), indicating that lower vitamin D levels were associated with longer and more febrile courses.

Multivariate linear regression analysis, adjusting for sex, diagnosis, and G-CSF use, showed that 25OHD level was not an independent predictor of neutropenia duration (Beta=-0.289, 95% CI: -0.706 to 0.127, p=0.171) or number of febrile days (Beta=-0.047, 95% CI: -0.112 to 0.019, p=0.160). Leukemia diagnosis was the only significant independent predictor for both outcomes (neutropenia duration: Beta=14.538, 95% CI: 7.603 to 21.474, p<0.001; febrile days: Beta=2.341, 95% CI: 1.251 to 3.430, p<0.001) (Table 3).

Discussion

This retrospective study examined the association between 25OHD levels and febrile neutropenia in pediatric cancer patients. Lower vitamin D levels were associated with longer neutropenia duration and more febrile days in univariate analysis. However, multivariate regression revealed that leukemia diagnosis, rather than vitamin D status, was the independent predictor of these outcomes. From a clinical perspective, patients with 25OHD levels below 15 ng/ml had a mean neutropenia duration of 22.4 days and 5.1 febrile days, compared to 16.4 days and 3.9 days in those with sufficient levels. This difference of approximately 6 days of neutropenia and 1.2 febrile days may translate into prolonged hospitalization, increased antibiotic use, and higher treatment costs, even if the association was not independent in multivariate analysis.

Our univariate findings are consistent with those of Gupta et al.[9], who reported that low vitamin D levels were associated with prolonged febrile neutropenia in pediatric patients. A large retrospective study by Penkert et al.[10] involving 378 children with newly diagnosed ALL showed that vitamin D deficiency at diagnosis was associated with a 1.7-fold increased risk of febrile neutropenia and a 1.73-fold increased risk of clinically documented infection during induction therapy.

Vitamin D exerts immunological effects through its nuclear receptor (VDR), expressed on T cells, B cells, macrophages, and dendritic cells. VDR activation induces antimicrobial peptides such as cathelicidin, enhances phagocytosis, and modulates cytokine responses. These mechanisms may be particularly relevant in neutropenic cancer patients [1,2]. García-García et al.[11] reported that cathelicidin—an antimicrobial peptide induced by vitamin D—was independently linked to sepsis progression in pediatric cancer patients with febrile neutropenia, supporting a potential mechanistic link.

The fact that 25OHD level lost its significance after adjusting for leukemia diagnosis suggests that the association between vitamin D and febrile neutropenia outcomes may be confounded by disease type. Leukemia patients receive more intensive and prolonged myelosuppressive chemotherapy, which independently

prolongs neutropenia and increases febrile episodes. This finding highlights the importance of controlling for diagnosis when interpreting vitamin D-outcome associations in pediatric oncology.

Vitamin D deficiency was found in 63.6% of our patients, consistent with the literature. Risk factors include reduced sun exposure during hospitalization, chemotherapy-related malabsorption, and corticosteroid use [6]. Patients who received vitamin D prophylaxis in infancy had significantly higher 25OHD levels (27.79±7.39 vs 13.16±6.36 ng/ml; $p<0.001$), underscoring the long-lasting benefit of early supplementation.

This study has several limitations. The retrospective, single-center design carries risk of bias and limits generalizability. The patient cohort was heterogeneous regarding malignancy type and chemotherapy protocols. Vitamin D levels were measured only at diagnosis, not during febrile episodes, which may not reflect levels at the time of neutropenia. Important confounders such as chemotherapy intensity, nutritional status, and corticosteroid use were not included in the regression model. The sample size was relatively small. Future prospective studies with larger, homogeneous cohorts and serial vitamin D measurements are needed.

Conclusion

Lower 25OHD levels were associated with longer neutropenia duration and more febrile days in pediatric cancer patients. However, this association was not independent after adjusting for leukemia diagnosis. Regular monitoring of 25OHD levels and supplementation in cases of deficiency may be clinically relevant. Further prospective studies are needed to clarify the independent role of vitamin D in this population.

Author contributions

Conception and design: B.O., G.Ş.; Data acquisition: B.O., G.Ş.; Data analysis: B.O., G.Ş.; Data interpretation: B.O., G.Ş.; Drafting of the manuscript: B.O.; Critical revision of the manuscript: B.O., G.Ş. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Health Sciences University Ankara Dr. Sami Ulus Research and Education Hospital Ethics Committee (Date: 09.07.2018, Decision/Protocol No: 2018-100). Written informed consent was not required due to the retrospective design of the study, as approved by the ethics committee.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that during the preparation of this study, the following AI-assisted technology was used: Claude (Anthropic) in May 2026. Extent of Use: Claude was used for language editing, formatting the manuscript and title page according to the journal's requirements, and drafting routine correspondence with the editorial office. All AI-assisted content was reviewed and edited by the authors, who take full responsibility for the accuracy and integrity of the manuscript. The authors confirm that they have critically reviewed and edited any AI-generated content and take full responsibility for the integrity, accuracy, and originality of the publication. The authors certify that the original human contribution is maintained and that AI-assisted tools are not listed or cited as authors.

References

[1] Mora JR, Iwata M, von Andrian UH. Vitamin effects on the immune system: vitamins A and D take centre stage. Nat Rev Immunol 2008;8(9):685-98. **[Crossref]**

[2] Holick MF. Vitamin D deficiency. N Engl J Med 2007;357(3):266-81. **[Crossref]**

[3] Garland CF, Garland FC, Gorham ED, et al. The role of vitamin D in cancer prevention. Am J Public Health 2006;96(2):252-61. **[Crossref]**

[4] Ginde AA, Mansbach JM, Camargo CA Jr. Vitamin D, respiratory infections, and asthma. Curr Allergy Asthma Rep 2009;9(1):81-7. **[Crossref]**

[5] Marti FM, Cullen MH, Roila F; ESMO Guidelines Working Group. Management of febrile neutropenia: ESMO clinical recommendations. Ann Oncol 2009;20(Suppl 4):166-9. **[Crossref]**

[6] Jackmann N, Mäkitie O, Harila-Saari A, Gustafsson J, Nezirevic Dernroth D, Frisk P. Vitamin D status in children with leukemia, its predictors, and association with outcome. Pediatr Blood Cancer 2020;67(4):e28163. **[Crossref]**

[7] Naeije L, de Silva MP, Hofman P. Can screening for low vitamin D levels prevent bone health complications in paediatric oncology patients? Cancer Rep (Hoboken) 2022;5(7):e1534. **[Crossref]**

[8] Misra M, Pacaud D, Petryk A, Collett-Solberg PF, Kappy M; Drug and Therapeutics Committee of the Lawson Wilkins Pediatric Endocrine Society. Vitamin D deficiency in children and its management: review of current knowledge and recommendations. Pediatrics 2008;122(2):398-417. **[Crossref]**

[9] Gupta V, Kumar V, Singh SK. Low vitamin D levels are associated with an adverse clinical outcome in febrile neutropenia. J Pediatr Hematol Oncol 2016;38(3):202-4. **[Crossref]**

[10] Penkert RR, Jones BG, Tang L, et al. Association of vitamin A and D deficiencies with infectious outcomes in children undergoing intensive induction therapy for acute lymphoblastic leukemia. J Pediatr 2024;273:114148. **[Crossref]**

[11] García-García PE, Palomo-Colli MA, Silva-Jivaja KM, et al. Cathelicidin, but not vitamin D, is associated independently with sepsis in pediatric patients with cancer and febrile neutropenia. Mol Clin Oncol 2024;22(2):22. **[Crossref]**

Acta Medica 2026; 57(3): 193-201 DOI: 10.32552/actamedica.2026.1204

acta medica

RESEARCH ARTICLE

Evaluating the proficiency of ChatGPT-3.5, ChatGPT-4o and Google Gemini in responding to frequently asked questions on gender affirming care by transgender people

Selin Tekin[1], Banu Ertürk[1], Seda Hanife Oğuz[1], Bülent Okan Yıldız[1]

[1]Division of Endocrinology and Metabolism, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Abstract

Background: Transgender and gender-diverse individuals encounter stigma, discrimination, and healthcare access barriers, leading them to seek online information, including from large language models like ChatGPT and Google Gemini. This study compares their responses on gender-affirming care.

Methods: This cross-sectional study evaluated responses from ChatGPT-3.5, ChatGPT-4o, and Google Gemini to 18 commonly asked questions on gender-affirming care, derived from current guidelines and clinical practice inquiries. Two endocrinologists independently assessed guideline adherence, while response reliability and quality were evaluated using the modified DISCERN scale and the Global Quality Scale, respectively. Hallucination tendency was rated using a Likert scale, and readability was analyzed using the Flesch Reading Ease, Flesch-Kincaid Grade Level, and Gunning Fog Index.

Results: The highest guideline compatibility was observed with ChatGPT-4o [81.1%], followed by ChatGPT-3.5 [77.2%] and Google Gemini [63.8%]. For quality ChatGPT-4o scored highest [4.8±0.3], ChatGPT-3.5 slightly lower [4.5±0.5], and Gemini significantly lower [2.4±0.5]. ChatGPT-4o and ChatGPT-3.5 showed similar quality [p=0.19] but both outperformed Gemini [p<0.001]. ChatGPT-4o had the greatest tendency to hallucinate. Overall, a high school education was needed to understand the responses. In terms of readability, Gemini's answers were easiest, followed by ChatGPT-3.5 and ChatGPT-4o [p<0.001].

Conclusions: ChatGPT-4o demonstrated the highest accuracy, quality, and guideline adherence in gender-affirming care, outperforming ChatGPT-3.5 and Google Gemini. However, its greater hallucination tendency necessitates caution in medical use. While Google Gemini had better readability, its lower compatibility and quality scores reduce its effectiveness in transgender care. These inaccuracies highlight that AI should not replace expert healthcare guidance.

Keywords: artificial intelligence, ChatGPT, Gemini, gender affirming care, transgender, endocrinology

Introduction

Transgender and gender diverse people encounter significant levels of stigma and discrimination in their daily lives compared to cisgender individuals [1,2]. They may find it challenging to open up to peers or parents, even doctors, or they may have already faced negative reactions when attempting to do so [2]. Transgender individuals face several challenges when seeking gender-affirming healthcare, including

Corresponding author: Melih Oral • **Email:** melihoral12@gmail.com

Received: October 20, 2025 **Accepted:** July 16, 2026 **Published:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

concerns related to hormone therapy, discrimination, and confidentiality [3]. Furthermore, many medical providers have received limited or no training on how to effectively work with transgender and gender diverse individuals [4], which may lead transgender people to question whether they will receive competent care. Consequently, trans and gender diverse individuals are known to avoid healthcare services due to concerns about encountering difficulties in reaching specialized professionals in transgender health because of the scarcity of healthcare providers experienced in this field [5]. Moreover, discrimination in the society may lead to unemployment or low income, making it harder to access healthcare [2].

Individuals facing limited access to healthcare services often turn to online sources for health information. Transgender and gender diverse individuals, in particular, demonstrate a higher tendency to search for health information online compared to cisgender individuals [6]. Large language models [LLMs] like ChatGPT and Google Gemini are a specialized type of artificial intelligence [AI] algorithm crafted for natural language processing tasks, with a central focus on predicting word sequences by taking into account the surrounding context. ChatGPT is an AI language model developed by OpenAI, based on the generative pre-trained transformer architecture [7]. ChatGPT is pre-trained on an extensive dataset that encompasses data from various online sources, such as websites, articles, books, and other publicly available texts [8]. Google Gemini is a family of large language models [LLMs] developed by Google DeepMind and open-sourced in December 2023. These models are multimodal, capable of understanding and processing information from various formats, including text, code, images, audio, and video [9]. Gemini models can be accessed through a Google account. The accessibility and open-access nature of these LLMs render them an appealing source of information for both patients and healthcare professionals. However, it is crucial not to underestimate current limitations, such as the absence of appropriate references; and to check whether the information is up-to-date, accurate, free of hallucinations, and unbiased [10].

ChatGPT and Google Gemini along with other similarly trained AI systems are subject to potential biases and dissemination of harmful misinformation, including stereotypes, gender norms, harmful dialogue, and deleterious opinions [11]. Transgender individuals may already be marginalized and stigmatized in various aspects of life and they are susceptible to systemic bias [12,13]. LLMs' comprehension of gender-affirming care and gender identity may be influenced by bias and misinformation, contingent upon factors such as the platform's training data, source code, and reinforced learning experiences. An AI system therefore require examination to ascertain its perspectives and accuracy regarding gender-affirming medical care, in order to evaluate its suitability as an educational resource.

AI systems, despite all these considerations, have nevertheless been widely used to search for answers to numerous medical questions. We considered that transgender individuals could also seek answers to their questions on ChatGPT or Google Gemini due to various reasons such as discrimination, stigma, and difficulties accessing medical services. Thus, we aimed to compare responses of ChatGPT-3.5, ChatGPT-4o and Google Gemini to common questions asked by transgender individuals with existing guideline information.

Methods

Study design

This is a cross-sectional non-human subject study, posing 18 commonly asked questions related to gender affirming care of transgender individuals to ChatGPT-3.5, ChatGPT-4o and Google Gemini. The questions were developed based on current guidelines [14,15], and the inquiries commonly asked by transgender individuals to us as experienced endocrinologists during our endocrinology outpatient clinical practice with transgender individuals.

The questions were categorized into three sections (Table 1);

1. Gender-affirming hormonal therapy [GAHT, n=12]
2. Adverse outcomes and long-term care [n=4]
3. Gender-affirming surgery [n=2]

A prompting question was posed first as 'Please answer the following questions on gender affirming care of trangender persons.' Each response from ChatGPT-3.5, ChatGPT-4o and Google Gemini were independently evaluated by two endocrinologists [S.T. and S.H.O.] using a 5-point Likert scale, based on adherence to the

Table 1. Compatibility of AI-generated answers to common questions about gender affirming care with current guidelines [5-point likert scale].

	ChatGPT-3.5		ChatGPT-4o		Google Gemini	
	ES	WP	ES	WP	ES	WP
A- Gender Affirming Hormonal Therapy [GAHT]						
What is gender affirming hormonal therapy [GAHT]?	5	5	5	5	5	5
What are the effects of GAHT?	4	4	5	5	3	3
Who are eligible for GAHT?	5	3	5	4	3	3
Which medications are used in GAHT, and how are they administered?	4	4	5	5	2	2
When do the effects of GAHT typically appear?	3	3	4	4	2	2
How long after starting GAHT does menstruation stop?	5	5	5	5	4	4
What additional steps can I take, besides GAHT, to promote facial hair growth?	2	2	2	2	2	2
When does the voice begin to change with GAHT, and how can I modify my voice?	NA	2	NA	5	NA	5
Does GAHT cause weight gain?	5	5	5	5	5	5
Is it possible to have children after starting GAHT?	3	3	4	4	2	2
Is GAHT a lifelong treatment?	NA	NA	NA	NA	NA	NA
Are treatments related to gender transition covered by health insurance?	NA	5	NA	5	NA	5
B-Adverse Outcomes & Long-Term Care						
What are the health risks and potential side effects of GAHT?	4	4	4	4	3	3
Do GAHT increase the risk of cancer?	5	5	5	5	3	3
How often should I be checked for my hormone levels and potential side effects?	5	5	5	5	3	3
Which blood tests are required to monitor GAHT?	5	5	5	5	4	4
C- Gender-Affirming Surgery						
Who are eligible for gender-affirming surgery?	5	5	5	5	3	3
What options are available for gender-affirming surgeries?	3	4	4	5	3	2

ES: Endocrine Society Guideline 2017 for endocrine treatment of gender dysphoric/gender incongruent persons [15], WPATH: World Professional Association for Transgender Health [WPATH] Standards of Care 8 [14], NA: Not available.

Endocrine Society [ES] Guideline 2017 for endocrine treatment of gender dysphoric/gender incongruent persons [15] and World Professional Association for Transgender Health [WPATH] Standards of Care 8 [14] (Figure 1):

1. Fully adheres to the guidelines without deviation
2. Mostly adheres but with minor omissions or deviations
3. Partially adheres with some notable gaps or inconsistencies
4. Minimally adheres and includes significant inaccuracies or gaps
5. Does not adhere to the guidelines or contradicts them

Evaluation of reliability, quality, tendency to hallucinate and readability

Each response from ChatGPT-3.5, ChatGPT-4o, and Google Gemini was independently evaluated by two endocrinologists. Reliability, quality, and tendency to hallucinate were scored by both evaluators, and a consensus score was determined (Figure 2).

The DISCERN scale, a widely used tool for assessing the reliability and quality of online health information, was employed in this study [16,17]. The scale is divided into three sections: the first section consists of eight questions assessing the reliability of the information, the second section includes seven questions evaluating the quality of information on treatment options, and the final section addresses the overall quality of the publication as a source of treatment information. This study utilized the modified DISCERN [mDISCERN] scale, which focuses exclusively on the first section of the original DISCERN scale. Responses were scored as follows: a "no" response was given a score of 1, a "partial" response was scored between 2 and 4, and a "yes" response received a score of 5. The total scores were categorized as follows: below 40% [8–15] was considered poor, 40–79% [16–31] was rated as fair, and above 80% [32–40] was deemed good [17].

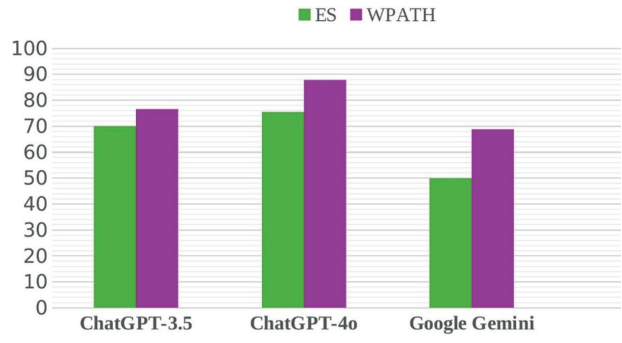


Figure 1. Compatibility of AI-generated responses with current guidelines.

ES: Endocrine Society Guideline 2017 for endocrine treatment of gender dysphoric/gender incongruent persons WPATH: World Professional Association for Transgender Health Standards of Care 8.

The Global Quality Scale [GQS], a tool commonly used in related studies, was applied to assess the quality of ChatGPT's responses [18]. On this scale, a score of 1 indicates poor quality, while a score of 5 represents excellent quality. The GQS also facilitates quality classification, with scores of 1–2 categorized as low quality, 3 as moderate quality, and 4–5 as high quality. ChatGPT has gained recognition as a remarkable global innovation, capable of producing highly realistic texts within seconds. However, it carries the potential risk of spreading inaccurate information and misconceptions, a phenomenon referred to by technical experts as "hallucination"[19]. Tendency to hallucinate of these responses were independently evaluated by two endocrinologists using a 5-point Likert scale.

Readability was assessed using Readable software [Readable.com, Horsham, United Kingdom] [20]. The analysis was conducted based on three widely recognized metrics: the Flesch Reading Ease [FRE] score, the Flesch-Kincaid Grade Level [FKGL], and the Gunning Fog Index [GFI]. An increase in the FRE score and a decrease in the other two show superior readability.

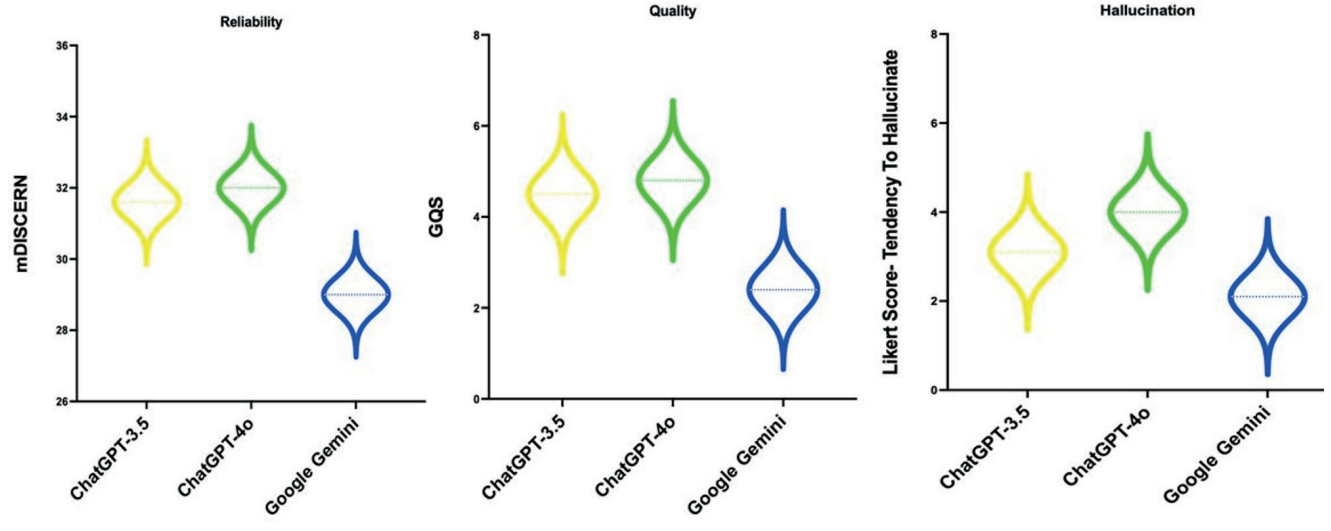


Figure 2. A. Reliability- B. quality- C. hallucination of ai genereted responses.

Data analysis

The Kolmogorov-Smirnov test was utilized to evaluate the normality of the data distribution. Numerical variables were presented as mean or median values. The accuracy and quality of the AI models were analyzed using frequencies and percentages and compared using Chi-square tests. For comparisons between chatbots, the mean values of accuracy, quality scores, and readability metrics were calculated and analyzed using the one-way ANOVA method. All statistical analyses were conducted using SPSS version 26.0 [IBM Corp., Armonk, NY, USA]. A p-value of <0.05 was considered the threshold for statistical significance.

Results

A total of 18 questions were posed to ChatGPT-3.5, ChatGPT-4o and Google Gemini regarding gender affirming care. Table 1 presents the compatibility of each answer of ChatGPT-3.5, ChatGPT-4o and Google Gemini with the Endocrine Society Guideline 2017 and WPATH Standards of Care 8 [14,15]. The overall compatibility of the responses from ChatGPT-3.5, ChatGPT-4o, and Google Gemini with the Endocrine Society and WPATH guidelines was 70% and 76.6%, 75.5% and 87.8%, and 50% and 68.8%; respectively. Responses to questions in Category B, which addressed adverse outcomes and long-term care, demonstrated the highest alignment with established guidelines. ChatGPT-3.5 and ChatGPT-4 achieved a 95% concordance with ES and WPATH guidelines, while Google Gemini achieved a 65% concordance with ES and WPATH. This was followed by responses in Category C, addressing gender-affirming surgery, where ChatGPT-4 achieved concordance rates of 90% with ES and 100% with WPATH guidelines. ChatGPT-3.5 demonstrated 80% and 90% concordance with ES and WPATH guidelines, respectively, while Google Gemini showed lower rates, at 60% for ES and 50% for WPATH. In contrast, responses in Category A, focused on gender-affirming hormone therapy [GAHT], demonstrated the lowest alignment with guidelines. ChatGPT-4 achieved concordance rates of 66.6% with ES and 81.6% with WPATH, while ChatGPT-3.5 showed 60% concordance with both. Google Gemini had the lowest alignment, with 46.6% to ES and 63.3% to WPATH (Table 1). The mean mDISCERN score, representing the "accuracy" of responses, was highest for ChatGPT-4o [32±0.4], followed by ChatGPT-3.5 [31.6 ± 2.4] and Google Gemini [29 ± 0.2] (Table 2). While ChatGPT-4o exhibited comparable accuracy to ChatGPT-3.5 [p = 0.12], it significantly outperformed Google Gemini [p < 0.001]. ChatGPT-3.5 also demonstrated significantly higher accuracy than Google Gemini [p < 0.001]. The mean GQS score, reflecting the "quality" of responses, was highest in ChatGPT-4o [4.8±0.3], followed by ChatGPT-3.5 [4.5±0.5] and Google Gemini [2.4±0.5]. ChatGPT-4o showed similar quality to ChatGPT-3.5 [p = 0.19] but significantly outperformed Google Gemini [p <0.001]. Additionally, ChatGPT-3.5 demonstrated higher quality compared to Google Gemini [p <0.001]. On the other hand, mean scores for 'tendency to hallucinate' was highest for ChatGPT-4o [4±0.7], followed by ChatGPT-3.5 [3.1±0.3] and Google Gemini [2.1±0.3], [p <0.001], Table 2. Among 'readability' indexes, Flesch–Kincaid grade level [FKGL] was highest for ChatGPT-4o [12.6±1.3], followed by ChatGPT-3.5 [12.1±1.8] and Google Gemini [11.8±1.6]. While FKGL did not differ between ChatGPT-4o and ChatGPT-3.5 [p = 0.95], ChatGPT-4o generated significantly more

Table 2. The accuracy, quality, tendency to hallucinate, and readability scores of AI-generated answers.

	ChatGPT-3.5	ChatGPT- 4o	Google Gemini
mDISCERN [mean±SD]	31.6±2.4	32 ± 0.4	29±0.2
GQS [mean±SD]	4.5±0.5	4.8±0.3	2.4±0.5
Tendency to hallucinate [mean±SD]	3.1±0.3	4±0.7	2.1±0.3
Readibility			
Flesch Kincaid Grade [mean±SD]	12.1±1.8	12.6±1.3	11.8±1.6
Gunning Fox Index [mean±SD]	12.8±1.8	13.1±1.8	13.4±2.4
Flesch Reading Ease [meanSD]	27±3.5	21.4±9.5	32.2±9.6

mDISCERN: modified DISCERN, SD: Standart Deviation, GQS: The Global Quality Scale .

readable responses than Google Gemini [lower FKGL; p < 0.001]. Gunning Fog Index [GFI] was similar between Google Gemini [13.4±2.4] and ChatGPT-4o [13.1±1.8] but exceeded that of ChatGPT-3.5 [12.8 ± 1.8] with statistical significance, [p<0.001]. Flesch Reading Ease [FRE] was the highest for Google Gemini [32.2±9.6] followed by ChatGPT-3.5 [27±3.5] and ChatGPT-4o [21.4±9.5] respectively, [p<0.001], Table 2.

Discussion

This study evaluated the efficacy and alignment of responses from ChatGPT-3.5, ChatGPT-4o, and Google Gemini with two current guidelines on the healthcare of transgender people [12,13]. The overall accuracy of the responses from ChatGPT-3.5, ChatGPT-4o, and Google Gemini with ES and WPATH guidelines was 70% and 76.6%, 75.5% and 87.8%, and 50% and 68.8%; respectively. The most notable differences in alignment with the two guidelines were observed in areas such as GAHT eligibility, voice therapy [which is addressed only superficially in the ES guideline], surgical options, and insurance coverage. This discrepancy may simply reflect the fact that the WPATH SOC8 is more up-to-date than the ES guideline. The ES guideline might not have included certain topics due to limited evidence at the time of its publication. Furthermore, the WPATH SOC8 is multidisciplinary and intended for a wide range of healthcare providers, aiming to offer more comprehensive transgender health care. The ES guideline, on the other hand, primarily targets endocrinologists and may prioritize hormonal aspects over adjunct treatments.

Nonetheless, the responses of ChatGPT-3.5 and Google Gemini to the question regarding the timing of GAHT effects were equally misaligned with both guidelines. In contrast to the guidelines, they suggested earlier timetables for skin changes and later timelines for fat redistribution when GAHT was started. Additionally, they included emotional changes, energy shifts, and mood alterations as expected effects of GAHT, which were not extensively covered in the guidelines. ChatGPT-4o distinguished itself by referring to WPATH guidelines and delivering more detailed answers, although some inaccuracies were still noted in its timelines. Nonetheless, all AI systems advised individuals about the importance of GAHT use under the supervision of healthcare professionals, and emphasized the individualized nature of the treatment. Likewise, the answers about the frequency of hospital visits emphasized that such visits can be tailored to individual needs and circumstances.

Notably, ChatGPT-3.5 and ChatGPT-4o provided non-evidence-based options for facial hair growth and voice therapy, beyond what is included in the guidelines. For facial hair growth, they suggested additional interventions such as micro-needling, exfoliation, supplements, herbal oils, and avoiding shaving, none of which are supported by evidence. For voice training, working with online apps in addition to working with a speech pathologist was offered. They also addressed fertility with optimism, despite limited evidence, while emphasizing the need to discuss fertility preservation before starting GAHT. Google Gemini struggled to provide clear answers to these questions, leading to shorter and less detailed responses. Gemini's responses were more generic, prioritizing consultation with healthcare professionals and referencing sources. On the other hand, it was not possible to assess the LLMs' alignment on the question of whether GAHT is a lifelong treatment, as this issue is not addressed in either guideline. ChatGPT-3.5 and -4o consider GAHT as a lifelong treatment for most transgender persons, depending on individual needs and circumstances.

Furthermore, it was observed that ChatGPT-3.5 and -4o mentioned surgical options beyond what is included in the guidelines, such as hip augmentation and muscle implants, which are not yet sufficiently covered in the literature. Both exhibited general knowledge and emphasized that the timing of GAS is a decision that can be influenced by eligibility, and many other factors such as individual preferences and the overall goals, as well as financial considerations. However, we did not pose other detailed questions about gender-affirming surgery in this study, as recent studies have already thoroughly assessed their responses related to this topic [21,22].

Regarding the accuracy and quality of answers related to gender-affirming care, Google Gemini's total score was significantly lower than that of ChatGPT-4o, while no significant difference was found between the scores of ChatGPT-4o and ChatGPT-3.5, consistent with findings from previous studies on various health topics [23-26]. The answers provided by Google Gemini and ChatGPT-3.5 were slightly better for topics related to adverse outcomes and gender-affirming surgery compared to GAHT. However, both were notably inferior to ChatGPT-4o across all areas. This difference can

be attributed to several factors. ChatGPT-3.5 was trained on data up to January 2022, while ChatGPT-4o, released in March 2023, is an improved version with a better model and more medical data. Additionally, ChatGPT-4o can fine-tune for specific domains and better understand specialized medical terminology in context. More importantly, ChatGPT-4o enhances the accuracy of answering questions in specific medical fields by building on user feedback from ChatGPT-3.5. Google Gemini, developed by Google is fundamentally different from ChatGPT-4o, developed by OpenAI, in how it processes information and generates answers [27-29]. However, consistent with previous studies, ChatGPT-4o demonstrated a higher tendency to produce hallucinations compared to ChatGPT-3.5 and Google Gemini [30,31].

Overall, it was noted that the readability level of all chatbots was fairly low. Based on the scores, only those with at least a high school education were able to understand the responses effectively. We determined that the readability order of AI answers regarding gender affirming care, from easy to difficult, is Gemini, and ChatGPT-3.5 and ChatGPT-4o which is consistent with a recent study [18,32,33].

While this study is among the first to explore the potential of AI in addressing transgender healthcare related answers, several limitations should be acknowledged. The lack of significant differences in response quality between ChatGPT-4 and 3.5 may be attributed to the relatively limited number of questions analyzed. Additionally, the subjective nature of response evaluation should be considered a methodological constraint, as individual interpretations may influence the assessment of accuracy and relevance. Another limitation is the study's inclusion of only two large language models, ChatGPT and Gemini, despite the availability of several alternatives, such as Perplexity.ai, Claude by Anthropic, Microsoft Copilot, Llama2 by Meta, and Mistral.ai. While these models may demonstrate superior performance in general AI applications, their effectiveness in medical and sensitive sociocultural contexts remains limited and largely unverified. The rapid advancement of AI technology indicates that chatbot performance is likely to evolve over time, potentially affecting the long-term applicability and generalizability of the study's findings. Additionally, despite the simultaneous administration of identical questions to the models, response variability at different time points remains a concern, as it may be influenced by updates to training datasets and modifications to model parameters.

In conclusion, while ChatGPT-4o demonstrated superior alignment with guidelines and provided more detailed responses compared to ChatGPT-3.5 and Google Gemini in the context of gender-affirming care, its tendency to produce hallucinations underscores the need for cautious application in medical contexts. Although AI models like ChatGPT and Google Gemini can serve as valuable tools for general medical information, they are not substitutes for professional medical advice. Patients should remain cautious when utilizing these AI models, as their responses may include inaccuracies or lack the specificity needed for individual situations. The variations in readability and accuracy among these AI models highlight the importance of continuous evaluation and refinement to ensure that they deliver accessible, reliable, and evidence-based information. This study emphasizes the necessity for healthcare professionals to verify AI-generated content and integrate these tools to support, rather than replace, clinical expertise. Therefore, patients are strongly advised to consult qualified healthcare providers to make informed decisions regarding their health and treatment options.

Author contributions

Conception: S.T., B.E., S.H.O., B.O.Y.; Design: S.T., S.H.O., B.O.Y.; Data acquisition: S.T., B.E.; Data analysis: S.T., B.E.; Data interpretation: S.T., B.E.; Drafting of the manuscript: S.T.; Critical revision of the manuscript: S.T., S.H.O., B.O.Y. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

Ethics committee approval and informed consent were not required for this study.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

References

[1] Spivey LA, Edwards-Leeper L. Future directions in affirmative psychological ınterventions with transgender children and adolescents. J Clin Child Adolesc Psychol 2019;48(2):343-56. **[Crossref]**

[2] Winter S, Diamond M, Green J, et al. Transgender people: health at the margins of society. Lancet 2016;388(10042):390-400. **[Crossref]**

[3] Call DC, Challa M, Telingator CJ. Providing affirmative care to transgender and gender diverse youth: disparities, ınterventions, and outcomes. Curr Psychiatry Rep 2021;23(6):33. **[Crossref]**

[4] Korpaisarn S, Safer JD. Gaps in transgender medical education among healthcare providers: a major barrier to care for transgender persons. Rev Endocr Metab Disord 2018;19(3):271-5. **[Crossref]**

[5] Teti M, Kerr S, Bauerband LA, Koegler E, Graves R. A Qualitative scoping review of transgender and gender non-conforming people's physical healthcare experiences and needs. Front Public Health 2021;9:598455. **[Crossref]**

[6] Heng A, Heal C, Banks J, Preston R. Transgender peoples' experiences and perspectives about general healthcare: a systematic review. International Journal of Transgenderism 2018;19(4):359-78. **[Crossref]**

[7] OpenAI. Introducing ChatGPT. Available at: https://openai.com/index/chatgpt/

[8] Floridi L, Chiriatti M. GPT-3: its nature, scope, limits, and consequences. Minds and Machines 2020;30:681-94. **[Crossref]**

[9] Google DeepMind. Gemini. Available at: https://deepmind.google/technologies/gemini/#introduction

[10] Wang C, Liu S, Yang H, Guo J, Wu Y, Liu J. Ethical considerations of using chatgpt in health care. J Med Internet Res 2023;25:e48009. **[Crossref]**

[11] Criss S, Nguyen TT, Gonzales SM, et al. "HIV Stigma Exists" - Exploring ChatGPT's HIV advice by race and ethnicity, sexual orientation, and gender identity. J Racial Ethn Health Disparities 2025;12(6):3622-35. **[Crossref]**

[12] Link BG, Phelan JC. Conceptualizing stigma. Annual Review of Sociology 2001;27(1):363-85. **[Crossref]**

[13] Velasco RAF, Slusser K, Coats H. Stigma and healthcare access among transgender and gender-diverse people: a qualitative meta-synthesis. J Adv Nurs 2022;78(10):3083-100. **[Crossref]**

[14] East Dundee I, Tax I, Vella B. World professional association for transgender health [WPATH].

[15] Hembree WC, Cohen-Kettenis PT, Gooren L, et al. Endocrine treatment of gender-dysphoric/gender-incongruent persons: an endocrine society clinical practice guideline. J Clin Endocrinol Metab 2017;102(11):3869-903. **[Crossref]**

[16] Ozduran E, Büyükçoban S. Evaluating the readability, quality and reliability of online patient education materials on post-covid pain. PeerJ 2022;10:e13686. **[Crossref]**

[17] Kumar VS, Subramani S, Veerapan S, Khan SA. Evaluation of online health information on clubfoot using the DISCERN tool. J Pediatr Orthop B 2014;23(2):135-8. **[Crossref]**

[18] Onder CE, Koc G, Gokbulut P, Taskaldiran I, Kuskonmaz SM. Evaluation of the reliability and readability of ChatGPT-4 responses regarding hypothyroidism during pregnancy. Sci Rep 2024;14(1):243. **[Crossref]**

[19] Ahmad Z, Kaiser W, Rahim S. Hallucinations in ChatGPT: an unreliable tool for learning. Rupkatha J Interdiscip Stud Humanit 2023;15(4):12. **[Crossref]**

[20] Readability Formulas. Avaliable at: https://readabilityformulas.com RcRsRtRIcRf

[21] Najafali D, Hinson C, Camacho JM, Galbraith LG, Tople TL, Eble D, et al. Artificial intelligence knowledge of evidence-based recommendations in gender affirmation surgery and gender identity: is ChatGPT aware of WPATH recommendations? Eur J Plast Surg 2023;46(6):1169-76. **[Crossref]**

[22] Snee I, Lava CX, Li KR, Corral GD. The utility of ChatGPT in gender-affirming mastectomy education. J Plast Reconstr Aesthet Surg 2024;99:432-5. **[Crossref]**

[23] Cinar C. Analyzing the performance of ChatGPT about osteoporosis. Cureus. 2023;15(9). **[Crossref]**

[24] Tong L, Zhang C, Liu R, Yang J, Sun Z. Comparative performance analysis of large language models: ChatGPT-3.5, ChatGPT-4 and Google Gemini in glucocorticoid-induced osteoporosis. J Orthop Surg Res 2024;19(1):574. **[Crossref]**

[25] Momenaei B, Wakabayashi T, Shahlaee A, et al. Appropriateness and readability of ChatGPT-4-Generated responses for surgical treatment of retinal diseases. Ophthalmol Retina 2023;7(10):862-8. **[Crossref]**

[26] Lim ZW, Pushpanathan K, Yew SME, et al. Benchmarking large language models' performances for myopia care: a comparative analysis of ChatGPT-3.5, ChatGPT-4.0, and Google Bard. EBioMedicine 2023;95. **[Crossref]**

[27] Mihalache A, Grad J, Patil NS, et al. Google Gemini and bard artificial intelligence chatbot performance in ophthalmology knowledge assessment. Eye (Lond) 2024;38(13):2530-5. **[Crossref]**

[28] Masalkhi M, Ong J, Waisberg E, Lee AG. Google DeepMind's gemini AI versus ChatGPT: a comparative analysis in ophthalmology. Eye (Lond) 2024;38(8):1412-7. **[Crossref]**

[29] Rane N, Choudhary S, Rane J. Gemini versus ChatGPT: applications, performance, architecture, capabilities, and implementation. SSRN Electronic Journal 2024. **[Crossref]**

[30] Aljamaan F, Temsah MH, Altamimi I, et al. Reference hallucination score for medical artificial ıntelligence chatbots: development and usability study. JMIR Med Inform 2024;12. **[Crossref]**

[31] Alkaissi H, McFarlane SI. Artificial hallucinations in ChatGPT: implications in scientific writing. Cureus 2023;15(2):e35179. **[Crossref]**

[32] Reyhan AH, Mutaf Ç, Uzun İ, Yüksekyayla F. A Performance evaluation of large language models in keratoconus: a comparative study of ChatGPT-3.5, ChatGPT-4.0, Gemini, Copilot, Chatsonic, and Perplexity. J Clin Med 2024;13(21):6512. **[Crossref]**

[33] McCarthy CJ, Berkowitz S, Ramalingam V, Ahmed M. Evaluation of an artificial ıntelligence chatbot for delivery of ır patient education material: a comparison with societal website content. J Vasc Interv Radiol 2023;34(10):1760-8 **[Crossref]**

Acta Medica 2026; 57(3): 202-210 DOI: 10.32552/actamedica.2026.1277

acta medica

RESEARCH ARTICLE

Clinical characteristics and legal outcomes of orthopedic malpractice claims in the constitutional court of Türkiye

Barış Kafa[1], Ulaş Can Kolaç[1], Mazlum Veysel Sili[1]

[1]Department of Orthopedics and Traumatology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Abstract

Objective: To evaluate the clinical characteristics, subspecialty distribution, and legal outcomes of orthopedic malpractice claims adjudicated through individual applications to the Turkish Constitutional Court, and to identify patterns of rights-based judicial review in orthopedic litigation.

Materials and Methods: A retrospective descriptive study was conducted using publicly accessible decisions retrieved from the official database of the Turkish Constitutional Court. A systematic search was performed using the keyword "ortopedi," and results were further refined by applying legal filters related to medical negligence and healthcare-associated harm. A total of 111 decisions were initially identified. After applying malpractice-related filters, 15 decisions underwent full-text evaluation. Of these, 13 cases met the predefined inclusion criteria and were included in the final analysis. Decisions were assessed for clinical relevance to orthopedic practice and reviewed for clinical context, orthopedic subspecialty involvement, primary malpractice allegations, alleged constitutional rights violations, admissibility status, and final judicial outcomes. Data were extracted using a standardized framework and analyzed descriptively using frequencies and percentages.

Results: Thirteen Constitutional Court decisions involving orthopedic care were included. Trauma-related management constituted the predominant clinical context, followed by postoperative complications and diagnostic delays. The most frequently involved subspecialties were trauma and fracture management, adult reconstructive surgery, and spine care. Constitutional violations were identified in 53.8% of cases (7/13), while the remaining 46.2% resulted either in findings of no violation or inadmissibility. Among cases reviewed on the merits, violations were primarily associated with procedural deficiencies in judicial evaluation, rather than direct determinations of medical fault.

Conclusion: Orthopedic malpractice claims reaching the Turkish Constitutional Court are primarily derived from trauma-related care. These findings should be interpreted within the framework of constitutional judicial review, which evaluates procedural adequacy and protection of fundamental rights rather than directly determining medical negligence. These results underscore the distinction between clinical malpractice litigation and constitutional judicial review in orthopedic disputes.

Keywords: orthopedic surgery, medical malpractice, trauma, constitutional court, patient rights

Corresponding author: Barış Kafa • **Email:** drkafa@gmail.com

Received: May 6, 2026 **Accepted:** February 23, 2026 **Published online:** August 14, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Introduction

Orthopedic surgery is a high volume medical specialty that includes trauma care, elective surgical procedures, and postoperative management [1]. Due to the technical complexity of treatment and the potential for serious complications, orthopedic practice has frequently been associated with malpractice claims across healthcare systems [2,3]. Fracture care, joint replacement surgery, and trauma-related procedures are among the most commonly reported clinical contexts in litigation studies [4-6].

A key issue in orthopedic malpractice cases is the distinction between unavoidable complications and deviations from accepted standards of care [3,7,8]. Adverse outcomes such as infection, neurovascular injury, thromboembolic events, and delayed bone healing may occur despite appropriate management; however, inadequate clinical assessment, delayed treatment, insufficient postoperative monitoring, or incomplete documentation may lead to allegations of negligence [1,7]. In this context, both clinical decision-making and the legal evaluation of medical conduct influence malpractice determinations [8,9].

Judicial review of medical malpractice claims relies largely on expert reports and the adequacy of judicial reasoning [8,10]. Limited causal analysis, insufficient consideration of clinical complexity, or inadequately justified court decisions may affect the effective protection of patient rights [8]. Higher courts serve a supervisory function in ensuring that fundamental rights, including the right to life and the protection of physical integrity, are appropriately upheld in healthcare-related disputes [3,8].

In Türkiye, the individual application mechanism to the Constitutional Court allows individuals to allege violations of constitutional rights arising from judicial proceedings related to medical malpractice [11]. Since 23 September 2012, individuals have been able to submit individual applications claiming violations of fundamental rights protected under the Constitution and the European Convention on Human Rights by public authorities [7,11]. Decisions of the Constitutional Court therefore provide valuable insight into both clinical allegations and the quality of lower court evaluations of medical evidence.

Although previous research has examined orthopedic malpractice using insurance claims and registry data, studies based on constitutional or supreme court decisions remain limited [5,12,13]. The association between clinical characteristics, orthopedic subspecialties, and constitutional-level legal outcomes has not been systematically investigated. While malpractice litigation is commonly evaluated through civil or criminal courts focusing on clinical fault, constitutional-level review primarily examines whether judicial processes adequately protected fundamental rights, representing a distinct layer of assessment.

This study aimed to analyze the clinical characteristics, subspecialty distribution, primary malpractice allegations, and legal outcomes of orthopedic malpractice claims reviewed in Constitutional Court decisions.

Materials and Methods

Study design

This study was designed as a retrospective descriptive analysis of orthopedic malpractice claims reviewed in decisions of the Constitutional Court of Türkiye. The search was conducted on 18.02.2026. Screening and eligibility assessment were performed by a single reviewer. The analysis focused on evaluating clinical characteristics, orthopedic subspecialties involved, primary malpractice allegations, and legal outcomes at the constitutional level. This study was based exclusively on publicly available judicial decisions and did not involve human participants or identifiable personal data. Therefore, institutional ethics committee approval was not required.

Data source and search strategy

All publicly available decisions were retrieved from the official Constitutional Court decision database (https://www.anayasa.gov.tr/tr/kararlar-bilgi-bankasi). A systematic search was performed using the keyword "ortopedi" (orthopedics). To ensure relevance to malpractice-related claims, results were further filtered using the legal categories "medical negligence/medical practices" and "death or harm resulting from medical

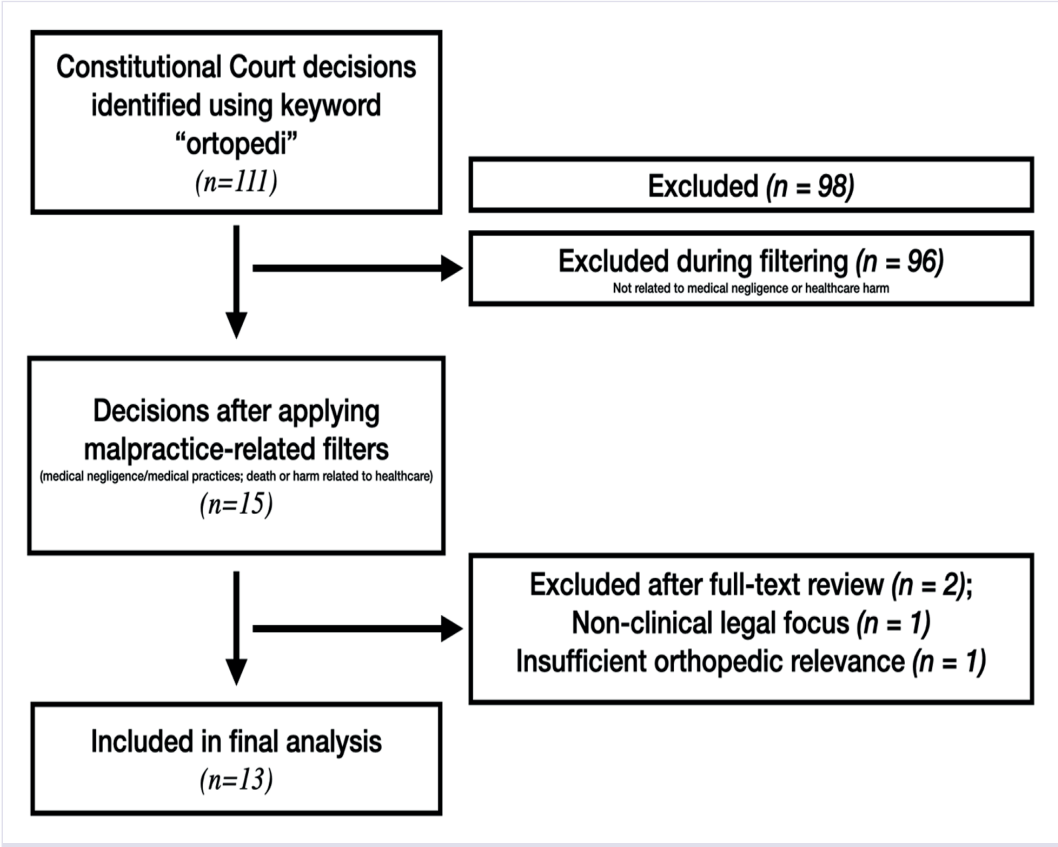


Figure 1. Flowchart illustrating the identification, screening, exclusion, and final inclusion of Constitutional Court decisions analyzed for orthopedic malpractice claims.

negligence or lack of access to healthcare services." The initial search yielded 111 decisions. After application of the malpractice-related filters, 15 decisions were identified for full-text review, Figure 1.

Eligibility criteria

Decisions were included if they involved orthopedic clinical care and contained allegations related to medical negligence, delayed diagnosis or treatment, inadequate management, or malpractice. Cases were excluded if they primarily addressed non-clinical issues, such as custodial care, ill treatment unrelated to medical practice, or procedural complaints without substantive evaluation of orthopedic care. Following full text review, two decisions were excluded due to insufficient clinical relevance, resulting in a final sample of 13 Constitutional Court decisions. Clinical relevance was defined as cases involving direct orthopedic diagnosis, treatment, or postoperative management.

Data extraction

For each included decision, clinical and legal data were systematically extracted from the full texts. Extracted variables included the clinical context of orthopedic care, the orthopedic subspecialty involved, the primary malpractice allegation, the reported clinical outcome, the Constitutional Court's ruling, and any remedies ordered.

Subspecialty classification

Orthopedic subspecialties were categorized according to the predominant clinical scenario described in each case, including trauma-related care, elective orthopedic surgery, postoperative complication management, pediatric orthopedic care, and hand surgery. When classification was uncertain, the clinical context was reassessed to ensure consistent categorization.

Outcome Measures and Analysis

The primary outcomes were the distribution of orthopedic subspecialties, types of malpractice allegations, and legal outcomes of Constitutional Court decisions. Secondary analysis examined clinical patterns associated with findings of constitutional violations. Data were analyzed descriptively using frequencies and percentages. Judicial reasoning was qualitatively reviewed to identify recurring factors contributing to constitutional violations.

Results

Trauma-related orthopedic care constituted the majority of included cases, accounting for 61.5% (8/13). Elective orthopedic surgery and postoperative complication management each represented 15.4% of cases (2/13), while pediatric and hand surgery contexts accounted for 7.7% (1/13) of cases (Table 1, Figure 2).

Delayed diagnosis or treatment was the most common malpractice allegation, occurring in 30.8% of cases (4/13). Inadequate clinical management accounted for 23.1% (3/13), followed by poor complication management and insufficient monitoring or follow-up, each representing 15.4% (2/13). Inadequate informed consent and multidisciplinary system-related delays were observed less frequently, each in one case (7.7%) (Table 2).

Constitutional violations were identified in 53.8% of cases (7/13). No violation was determined in 30.8% (4/13), while 15.4% (2/13) of applications were declared inadmissible (Figure 3).

Among trauma-related cases, constitutional violations were observed in 50.0% (4/8), whereas two cases resulted in no violation and two were deemed inadmissible. Both cases involving postoperative complication management resulted in findings of constitutional violations. In elective orthopedic surgery, one case resulted in a violation and one in no violation, while the single pediatric or hand surgery case resulted in no violation (Table 3). Importantly, constitutional violation findings were predominantly related to deficiencies in judicial reasoning, expert evaluation, or procedural safeguards rather than direct confirmation of orthopedic malpractice.

Discussion

This study demonstrates that trauma-related orthopedic care constituted the majority of cases reviewed at the constitutional level; however, these findings reflect the characteristics of disputes subjected to rights-based judicial review rather than the overall incidence of orthopedic malpractice. Delayed diagnosis or treatment and inadequate clinical management were the most frequently observed allegations. Cases involving postoperative complication management demonstrated a consistently high rate of constitutional violations.

The predominance of trauma related care among malpractice claims is consistent with the clinical characteristics of orthopedic practice [7,14]. Trauma management frequently occurs in emergency settings where rapid assessment and decision making are required [15]. Injury patterns may be complex, and early clinical findings can evolve over time [15]. These factors may increase the likelihood of diagnostic uncertainty, delayed recognition of complications, or incomplete documentation. Furthermore, trauma care often involves coordination among multiple medical teams, and communication gaps or referral delays may contribute to adverse outcomes [15,16].

Delayed diagnosis or treatment represented the most common allegation in the present series. In orthopedic trauma, early identification of fracture patterns, neurovascular compromise, or compartment syndrome is critical [17]. Even relatively short delays in appropriate intervention may lead to permanent functional impairment [18,19]. From a legal perspective, delays in care may facilitate the establishment of a causal relationship between clinical management and adverse outcomes [20]. As a result, such cases may be more likely to progress through judicial review.

An important finding was the consistent identification of constitutional violations in cases related to postoperative complication management [21]. Postoperative care requires systematic monitoring and timely response to clinical deterioration. Although complications may occur despite appropriate surgical technique, failure to recognize or manage these complications in a timely manner may be interpreted as a deviation from expected standards of care [22]. The findings suggest that courts may evaluate postoperative management with particular attention to the adequacy of monitoring and responsiveness.

Table 1. Clinical and legal characteristics of orthopedic malpractice cases reviewed by the Constitutional Court.

	Decision no.	Clinical context	Subspecialty	Primary allegation	Clinical outcome	Court decision	Remedy
1	2013/2803	Post-traumatic orthopedic care	Trauma	Inadequate fracture management	Permanent impairment	Inadmissible	None
2	2014/19349	Trauma-related orthopedic treatment	Trauma	Delayed or inadequate treatment	Severe injury	Inadmissible	None
3	2016/11550	Orthopedic treatment following injury	Trauma	Inadequate clinical management	Persistent disability	No violation	None
4	2017/18143	Postoperative complication	Postoperative care	Inadequate complication management	Neurological deficit	Violation	Retrial
5	2017/40500	Neonatal orthopedic injury	Pediatric orthopedics	Inadequate fracture and nerve injury management	Permanent functional loss	No violation	None
6	2019/4610	Trauma-related care	Trauma	Delayed diagnosis and treatment	Long-term disability	Violation	Retrial and compensation
7	2019/11247	Elective orthopedic surgery	Elective surgery	Inadequate informed consent	Permanent impairment	Violation	Retrial
8	2020/10663	Upper extremity fracture	Hand surgery	Inadequate immobilization	Delayed surgical intervention	No violation	None
9	2020/16941	Elective orthopedic surgery	Elective surgery	Alleged surgical mismanagement	Functional limitation	No violation	None
10	2020/22429	Postoperative mortality	Postoperative care	Failure to manage complications	Death	Violation	Retrial
11	2020/26359	Trauma-related orthopedic treatment	Trauma	Insufficient follow-up	Permanent disability	Violation	Retrial
12	2021/12621	Trauma management	Trauma	Insufficient clinical monitoring	Severe disability	Violation	Retrial
13	2021/51903	Major orthopedic trauma	Trauma	Delayed multidisciplinary intervention	Death	Inadmissible	None

The relatively high proportion of constitutional violations in this cohort warrants careful interpretation. Constitutional review does not primarily reassess medical fault but evaluates whether judicial proceedings adequately protected fundamental rights, including the right to life and bodily integrity [7,11]. In several decisions, violations were associated not only with clinical concerns but also with deficiencies in expert evaluation, judicial reasoning, or lower assessment [11]. This distinction highlights that constitutional findings primarily reflect procedural inadequacies in judicial evaluation rather than direct confirmation of medical negligence, emphasizing the need to interpret these results within their legal context.

The distribution of subspecialties and allegations observed in this study is broadly aligned with patterns reported in orthopedic malpractice literature, where

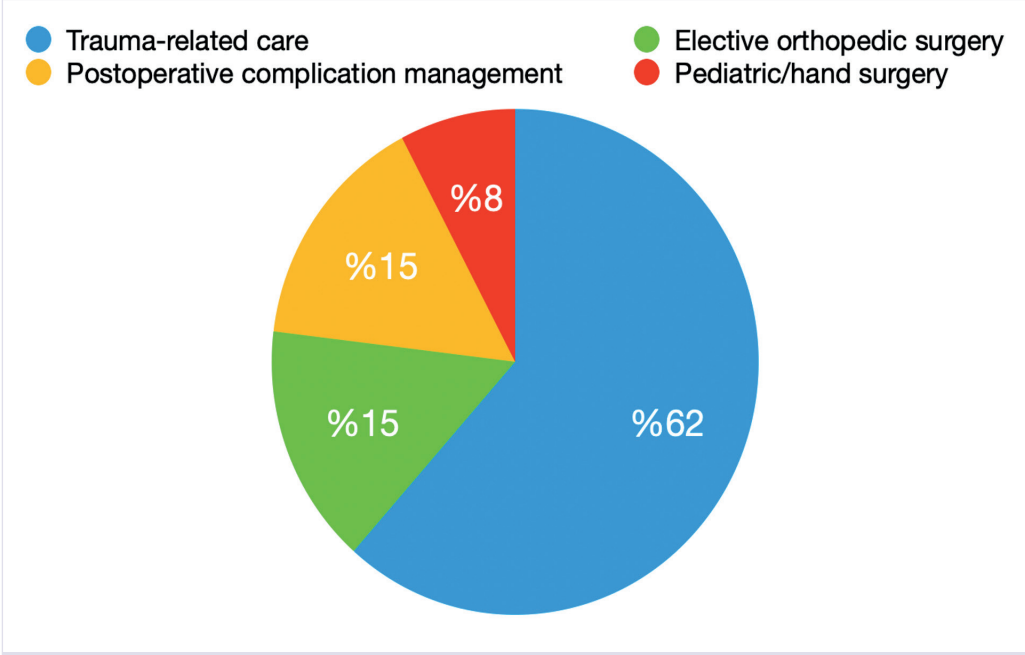


Figure 2. Distribution of orthopedic subspecialties involved in malpractice claims reviewed by the Constitutional Court.

Table 2. Distribution of primary malpractice allegations in orthopedic constitutional court cases.

Primary allegation	Number of cases	Percentage (%)
Delayed diagnosis or treatment	4	30.8
Inadequate clinical management	3	23.1
Poor complication management	2	15.4
Insufficient monitoring or follow-up	2	15.4
Inadequate informed consent	1	7.7
Multidisciplinary/system-related delay	1	7.7
Total	13	100

trauma care, fracture management, and delays in diagnosis are commonly identified risk areas [7,8]. Özdemir et al. examined 71 orthopedic malpractice cases from the Supreme Court and found that most claims arose from trauma care, diagnostic delays, and postoperative complications, with physicians ultimately found at fault in about one-third of cases, while many decisions were overturned due to inadequate expert evaluation [7]. In contrast, our Constitutional Court analysis of 13 cases shows a much lower rate of established rights violations, with the majority of applications dismissed as inadmissible or resulting in no violation, indicating that while clinical allegations are similar, only a small proportion escalate to constitutional level breaches of patient rights. These findings are also consistent with international malpractice literature, where trauma-related care, delayed diagnosis, and postoperative complications are among the most frequently reported causes of litigation [8,14]. However, unlike conventional malpractice studies that directly assess physician liability, the present study reflects the subset of cases escalating to constitutional review, where procedural adequacy and protection of fundamental rights play a central role.

While direct clinical recommendations cannot be derived from constitutional decisions, the findings highlight recurrent clinical contexts, such as trauma care, delayed

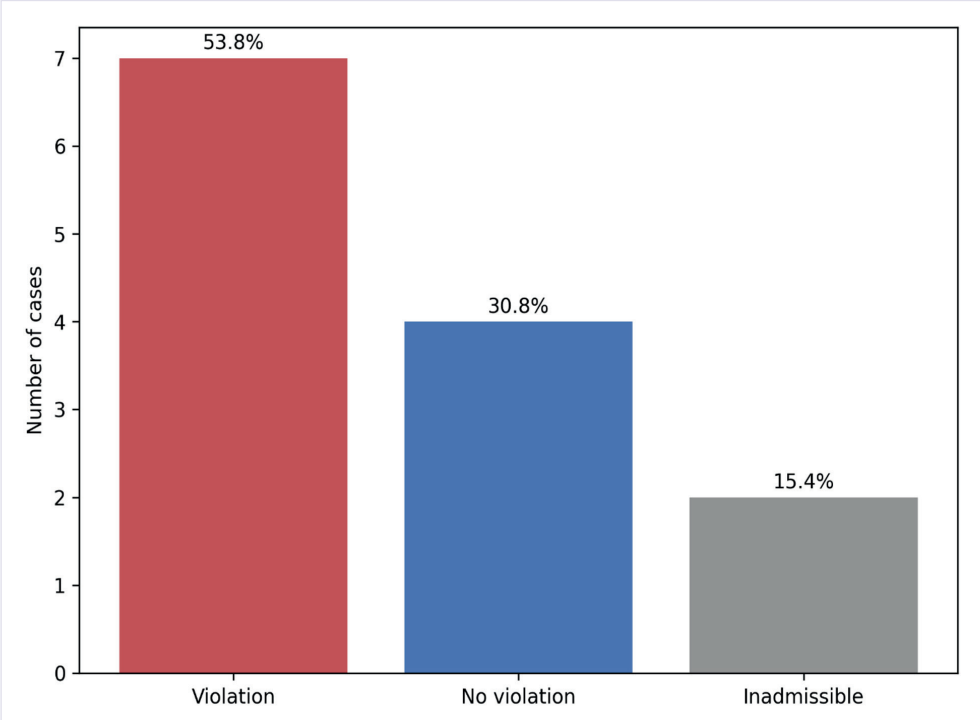


Figure 3. Legal outcomes of orthopedic malpractice claims adjudicated by the Constitutional Court.

Table 3. Orthopedic subspecialties and Constitutional Court legal outcomes.

Subspecialty	Violation	No violation	Inadmissible	Total
Trauma-related care	4	2	2	8
Elective orthopedic surgery	1	1	0	2
Postoperative complication management	2	0	0	2
Pediatric/hand surgery	0	1	0	1
Total	7	4	2	13

diagnosis, and postoperative management, that are frequently associated with medicolegal disputes. Structured evaluation protocols in trauma settings, clear documentation of decision-making processes, and timely referral to appropriate subspecialists may reduce the risk of adverse outcomes. In the postoperative period, standardized monitoring and early recognition of complications may be particularly relevant. Strengthening communication among multidisciplinary teams may also reduce system related delays.

Several limitations should be acknowledged. The number of cases is limited, reflecting the selective nature of Constitutional Court review rather than the overall frequency of orthopedic malpractice claims. The descriptive design does not allow statistical modeling. In addition, reliance on judicial decisions restricts access to detailed clinical data. Despite these limitations, the study shows the types of orthopedic cases that reach constitutional adjudication and the clinical circumstances most frequently associated with adverse legal outcomes.

Conclusion

In conclusion, orthopedic malpractice claims reviewed at the constitutional level most commonly involve trauma related care and are frequently associated with delayed diagnosis or treatment and inadequate clinical

management. Although constitutional violations were identified in a substantial proportion of cases, these findings often reflect procedural deficiencies in judicial evaluation rather than direct determinations of medical fault. Attention to timely clinical intervention, structured postoperative monitoring, and comprehensive documentation may contribute to improved patient outcomes and reduced legal exposure.

Author contributions

Conception: B.K.; Design: B.K.; Data acquisition: U.C.K., M.V.S.; Data analysis: U.C.K., M.V.S.; Drafting of the manuscript: U.C.K., M.V.S.; Critical revision of the manuscript: B.K. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

Ethics committee approval and informed consent were not required for this study.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

References

[1] Dun C, Rumalla KC, Walsh CM, Escobar C. Evaluating patient and surgeon characteristics associated with care cost and outcomes for knee and hip replacement procedures: a national medicare cohort study. JB JS Open Access 2024;9(1):e23.00088. **[Crossref]**

[2] Carroll AE, Buddenbaum JL. High and low-risk specialties experience with the U.S. medical malpractice system. BMC Health Serv Res 2013;13:465. **[Crossref]**

[3] Jena AB, Seabury S, Lakdawalla D, Chandra A. Malpractice risk according to physician specialty. N Engl J Med 2011;365(7):629-36. **[Crossref]**

[4] Kwan JL, Calder LA, Bowman CL, et al. Characteristics and contributing factors of diagnostic error in surgery: analysis of closed medico-legal cases and complaints in Canada. Can J Surg 2024;67(1):E58-E65. **[Crossref]**

[5] Sailhan F, Bouché PA, Delaunay C, Hamadouche M, Chatellard R. Reasons for malpractice claims after primary total hip arthroplasty in France: insurance data from 240 claims from 2014 to 2017. Orthop Traumatol Surg Res 2024;110(5):103885. **[Crossref]**

[6] Megalla M, Zaifman JM, Grace ZT, Imam N, Kohan EM, Alberta FG. Medical malpractice litigation following humeral fractures in the United States. Shoulder Elbow 2025;17(1):43-9. **[Crossref]**

[7] Özdemir U, Akar A, Serttaş MF, Başer A. Fifteen years of orthopedic malpractice litigation in türkiye: a supreme court analysis and international comparison. J Clin Med 2026;15(2):625. **[Crossref]**

[8] Gill SS, Sugand K, Gupte CM. Why do orthopaedic surgeons get sued? An analysis of £2.2 billion in claims against NHS England: trends in litigation and strategies to enhance care. Arch Orthop Trauma Surg 2025;145(1):351. **[Crossref]**

[9] A'Court J, Yassa R, Charalambous CP. Litigation related to casting in Orthopaedics-An analysis of claims against the National Health Service in England. Injury 2017;48(7):1405-7. **[Crossref]**

[10] Lane J, Bhome R, Somani B. National trends and cost of litigation in UK National Health Service (NHS): a specialty-specific analysis from the past decade. Scott Med J 2021;66(4):168-74. **[Crossref]**

[11] Necmiye ŞE, Mehmet A, Ömer K, Mustafa A. Tıbbi uygulama hatalarina yönelik türkiye cumhuriyeti anayasa mahkemesi kararlarinin incelenmesinde; anesteziyoloji ve reanimasyon uzmanliği ile ilgili hatali tibbi uygulama iddialarinin tespiti ve analizi: kesitsel araştirma. Türkiye Klinikleri Anesteziyoloji Reanimasyon Dergisi 2025;23(1):15-26. **[Crossref]**

[12] Latario LD, Deter CJ, Deter AM, Baratz ME. The increasing medicolegal cost in the care of high-level athletes. J Shoulder Elbow Surg 2025;34(1):414-9. **[Crossref]**

[13] Sönmez MM, Alpsoy MY, Özdemir E, Sevin A, Kaygusuz MA. Malpractice litigation after total hip arthroplasty: A legal 10-year database. Jt Dis Relat Surg 2023;34(2):389-95. **[Crossref]**

[14] Thabet AM, Adams A, Jeon S, et al. Malpractice lawsuits in orthopedic trauma surgery: a meta-analysis of the literature. OTA Int 2022;5(3):e199. **[Crossref]**

[15] Creager AE, Siebler JC. Optimizing the orthopaedic trauma patient- staged management. operative techniques in orthopaedics 2024;34(3):101127. **[Crossref]**

[16] Greco T, Bernasconi A, Perisano C. Trauma and orthopedic surgery: recent developments and future challenges. J Clin Med 2025;14(13):4654. **[Crossref]**

[17] Laasonen EM, Kivioja A. Delayed diagnosis of extremity injuries in patients with multiple injuries. J Trauma 1991;31(2):257-60.

[18] Crist BD, Ferguson T, Murtha YM, Lee MA. Surgical timing of treating injured extremities: an evolving concept of urgency. Instr Course Lect 2013;62:17-28.

[19] Gyftopoulos S, Chitkara M, Bencardino JT. Misses and errors in upper extremity trauma radiographs. AJR Am J Roentgenol 2014;203(3):477-91. **[Crossref]**

[20] Ravery AS, Villatte G, Dartus J, Descamps S, Boisgard S, Erivan R. Reasons for delays to orthopaedic and trauma surgery: a retrospective five-year cohort. Orthop Traumatol Surg Res 2023;109(4):103535. **[Crossref]**

[21] Saud AM, Raza T, Kamran M, et al. Complications and revision surgery in orthopedics focusing on traumatology. Cureus 2024;16(7):e65640. **[Crossref]**

[22] Ricketts D, Rogers RA, Roper T, Ge X. Recognising and dealing with complications in orthopaedic surgery. Ann R Coll Surg Engl 2017;99(3):185-8. **[Crossref]**

Acta Medica 2026; 1281(3): 211-219 DOI: 10.32552/actamedica.2026.

acta medica

RESEARCH ARTICLE

Quantitative CT assessment of thymic involution in adults: Role of age, sex and smoking

Selin Ardalı Düzgün[1], İlke Taşçı[1], Jale Karakaya[2], Gamze Durhan[1]

[1]Department of Radiology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
[2]Department of Biostatistics, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Abstract

Objective: Thymus size and density progressively decrease through adulthood on CT. Many factors may influence thymic involution, resulting in varying CT appearances. This study aimed to evaluate thymic involution in adults using a detailed quantitative CT analysis—including volumetric soft-tissue segmentation and density measurements—and to investigate its associations with age, sex, and smoking.

Materials and Methods: Adults aged 18-50 years without known systemic disease, who underwent chest CT between June 2020 and October 2025 were retrospectively reviewed. Visual thymic scoring (0-3) and quantitative CT measurements—transverse and anterior-posterior diameters, right and left lobe lengths, overall thymus area, soft tissue area, and attenuation—were performed. These parameters were compared among thymic score groups, also based on sex and smoking status. Thymic scores were dichotomized into advanced (>50% fatty replacement, scores 0-1) and mild (≤50%, scores 2-3) groups. Logistic regression analysis was performed to identify independent predictors of advanced fatty replacement.

Results: A total of 136 patients (mean age 35.5±9, 48.5% male) were included. Age and sex significantly differed across thymic score groups ($p<0.001$ and $p=0.018$, respectively). Transverse diameter, CT attenuation, and soft tissue area increased with higher thymic scores (all $p<0.05$), although other diameters were not significantly different between score groups. Males exhibited significantly lower thymic scores and CT attenuation despite higher thymus diameters and area ($p<0.05$). Age and male sex were independently associated with advanced fatty replacement ($p<0.05$). Smoking was not an independent predictor despite negative correlations with thymic score, CT attenuation, and soft tissue area.

Conclusion: Thymic involution in adults is heterogeneous and primarily driven by age and sex. Fatty replacement predominates over size reduction, and males are more likely to exhibit advanced involution.

Keywords: thymus, chest, multidetector computed tomography

Introduction

The thymus is a lymphoid organ located in the anterior mediastinum, responsible for T-cell development. Physiologically, the thymus reaches its maximum size in puberty, followed by a gradual involution with increasing age. The epithelial component regresses, and fatty replacement occurs [1-3]. Correspondingly, on computed tomography (CT), the thymus appears as a homogeneous soft-tissue structure, with thymic size and density progressively decreasing through adulthood. This involution process varies among individuals, resulting in different appearances and stages of thymic involution in CT even within the same

Corresponding author: Selin Ardalı Düzgün • **Email:** selin.ardali@gmail.com

Received: March 1, 2026 **Accepted:** March 31, 2026 **Published online:** August 14, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

age group [3-7]. Awareness of these differences and age-related changes in thymic size and density is important for distinguishing a normal thymus from thymic hyperplasia or thymic masses.

Physiological thymic involution is responsible for functional impairment of T-cell-mediated immunity and was associated with age-related immunologic decline, also known as "immunosenescence". Since aging and impaired immune function are associated with an increased incidence of infections and malignancies, the factors affecting, especially accelerating thymic involution, are important [8,9].

Previous studies have reported that age, sex, smoking, and body mass index are associated with the degree of thymic fatty involution on CT [3,4,10].

In this study, we aimed to evaluate thymic involution in adults using a detailed quantitative CT analysis in addition to visual assessment—including volumetric soft-tissue segmentation and density measurements—and to investigate its associations with age, sex, and smoking.

Material and Methods

Study population

Our institutional review board approved the study (SBA 25/987, approval number: 2025/22-36). Patients who underwent chest CT between June 2020 and October 2025 were retrospectively reviewed from our institutional database. Patients who met the following inclusion criteria were included in the study: patients (i) between 18 and 50 years of age, (ii) without a known systemic, oncologic, or immunological disease, (iii) without a known history of thoracic or mediastinal surgery, and (iv) with available data regarding smoking history.

Exclusion criteria were as follows: patients (i) younger than 18 or older than 50 years of age, (ii) with a known malignancy or systemic disease, (iii) with a history of chemotherapy, immunotherapy, radiotherapy, bone marrow transplantation or high-dose steroid treatment, (iv) with a history of thymic mass or evidence of a thymic mass on current CT, (v) who underwent thoracic or mediastinal surgery, (vi) with unavailable smoking data, and (vii) with poor-quality CT scans due to artifacts.

Patient characteristics such as age, sex, smoking status, and cumulative smoking amount (pack-years) were retrospectively retrieved from electronic medical records.

CT protocol

All patients underwent chest CT in the supine position, after intravenous contrast administration, during breath-hold at the end of inspiration. CT scans were obtained using multidetector CT scanners (Somatom Force and Somatom Perspective, Siemens Healthineers, Erlangen, Germany; and Optima CT 540, GE Healthcare, Chicago, IL, USA). Scanning parameters were as follows: tube voltage 100-120 kV, tube current 100-300 mAs (depending on patient size), matrix 512×512, field of view 350×350 mm, and slice thickness between 3 and 5 mm. Images were reconstructed using iterative reconstruction with a slice thickness between 1 and 1.5 mm and a soft-tissue kernel.

CT evaluation

CT images in mediastinal window settings (WW=350 HU, WL=50 HU) were analyzed by two thoracic radiologists with 6 and 10 years of experience using a dedicated workstation (MM Reading, syngo.via version VB60S_HF03, Siemens Healthineers, Erlangen, Germany). First, visual thymic scoring was performed as follows: score 0, complete fatty replacement; score 1, predominantly fatty tissue; score 2, approximately half of the thymic gland is composed of fatty tissue and half soft tissue; score 3, predominantly soft tissue [4] (Figure 1).

Following visual scoring, CT-based thymic measurements were performed in thymic score groups 1, 2, and 3. Patients with a thymic score of 0 were excluded from quantitative CT measurements due to complete fatty replacement. The axial slice demonstrating the largest thymic area was identified. Using this slice, the transverse diameter, anterior-posterior (AP) diameter, length, and thickness of the right and left thymic lobes were measured as described previously [3] (Figure 2).

The maximum area and mean attenuation (Hounsfield unit- HU) of the thymus were measured using a region-of-interest (ROI), avoiding mediastinal fat tissue. Semiautomatic volumetric segmentation of thymic soft tissue (restricted to soft tissue CT attenuation, between -50 HU and 150 HU) was performed using the "Region growing" tool (syngo.via version VB60S_HF03, Siemens

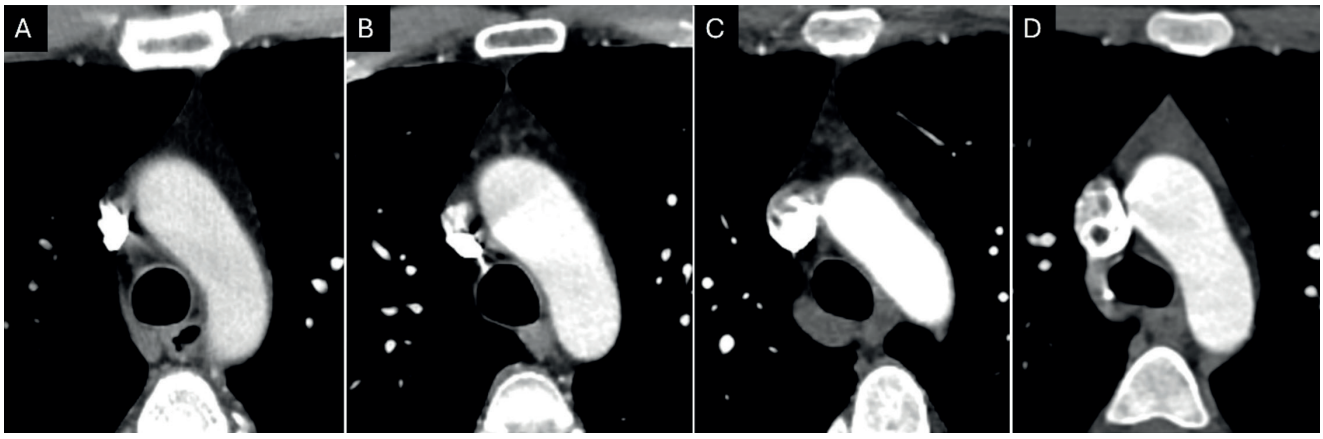


Figure 1. (A) Score 0: Complete fatty replacement in a 49-year-old male (B) Score 1: Predominantly fatty tissue in a 41-year-old male (C) Score 2: Approximately equal fatty and soft-tissue components in a 47-year-old female (D) Score 3: Solid thymic tissue in a 23-year-old female.

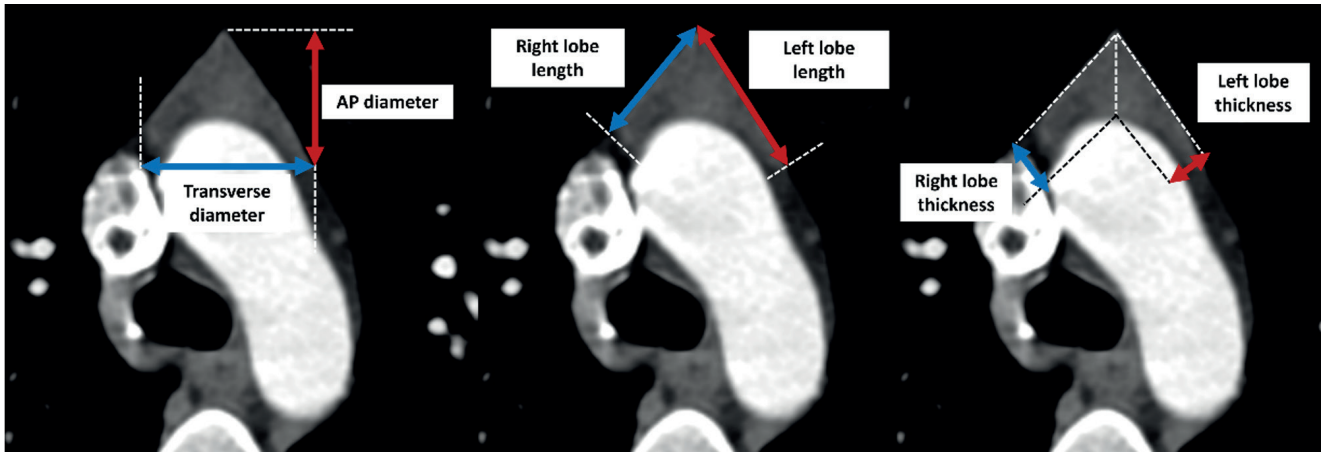


Figure 2. Measurements of thymic diameters on CT.

Healthineers, Erlangen, Germany). Subsequently, the thymic soft-tissue area was derived from volumetric data using the following formula: volume (cm3)/ [slice thickness (mm) /10] (Figure 3). Also, the soft tissue-to-overall thymus area ratio was calculated by dividing the thymic soft tissue area by the overall thymus area.

Statistical analysis

Normality of the data distribution was evaluated using the Shapiro-Wilk and Kolmogorov-Smirnov tests. Continuous variables were presented as mean ± standard deviation (SD) for normally distributed data, and as median (interquartile range) for non-normally distributed data. Categorical variables were presented as frequencies (percentages) and compared using the Pearson chi-square test.

For comparisons of continuous variables among different thymic score groups, one-way analysis of variance (ANOVA) was used for normally distributed data, while the Kruskal-Wallis test was used for non-normally distributed data. Post-hoc pairwise comparisons were conducted using Tukey's HSD following ANOVA, and Dunn's test with Bonferroni correction following the Kruskal-Wallis test. For numerical variables that show a normal distribution but have non-homogeneous variances, the Welch ANOVA test was used, and for

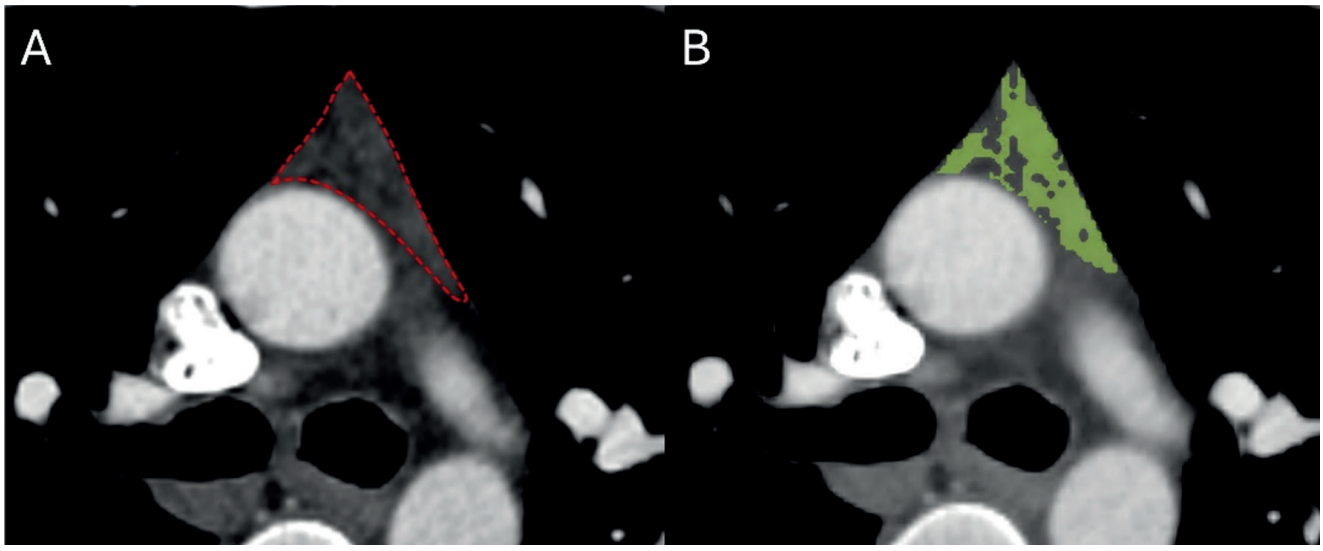


Figure 3. Measurement of overall thymic area (A) and soft-tissue area restricted to soft tissue attenuation between -50 HU and 150 HU (B).

variables showing differences after the test, the Tamhane post hoc test was used. To compare differences based on smoking status and sex groups, the Mann-Whitney U test was used.

Thymic scores were dichotomized into two groups based on the degree of fatty replacement: advanced fatty replacement (>50%, scores 0- 1), and mild fatty replacement (≤50%, scores 2-3). Logistic regression was used to investigate the predictors of advanced fatty replacement.

All analyses were performed using IBM SPSS Statistics 23.0 software. A p-value < 0.05 was considered statistically significant.

Results

A total of 136 patients (mean age 35.5±9, 48.5% male) were included in the study. The most common CT indications were chest pain (n=36, 26.5%), dyspnea (n=17, 12.5%), suspected X-ray abnormality (n=17, 12.5%), trauma (n=16, 11.8%), screening due to smoking history (n=15, 11%), donor evaluation (n=13, 9.6%), cough (n=12, 8.8%), screening for malignancy (n=6, 4.4%) and fever (n=4, 2.9%).

Patients were grouped according to thymic scores as follows: score 0 (n=14, 10.3%), score 1 (n=37, 27.2%), score 2 (n=37, 27.2%), and score 3 (n=48, 35.3%). Age and sex were significantly different between the thymic score groups (p<0.001 and p=0.018, respectively). Higher thymic scores were observed in younger patients; in contrast, lower thymic scores, corresponding to advanced fatty replacement, were associated with older age. While 69.8% of patients aged 18-30 years had a thymic score of 3, this ratio was 13% among those aged 41-50 years (Table 1).

The most common thymic scores among female and male patients were score 3 (n=31, 47%) and score 1 (n=22, 31.4%), respectively. Most patients with a score of 0 were male (n=11, 78.6%), while the majority of those with a score of 3 were female (n=31, 64.6%). Also, score 0 was the least observed thymic score among females (n=3, 4.5%) (Table 1).

Fifty-six percent of the patients had a smoking history (n=76). The majority of patients with a thymic score of 0 (n=9, 64.3%) had a smoking history, whereas more than half of the patients with a score of 3 had no smoking history (n=27, 56.3%). Although smoking status was not significantly different between groups, among non-smokers, score 3 was the most frequent thymic score (n=27, 45%). Cumulative smoking exposure was significantly different between groups. Notably, the

Table 1. Patient characteristics according to thymic score groups.

	Thymic score 0 (n=14)	Thymic score 1 (n=37)	Thymic score 2 (n=37)	Thymic score 3 (n=48)	p value	Post-hoc comparison
Age	43.9±5.7	41.5±5.7	34.5±8	29.3±8	<0.001[a]	0 vs 1, p=0.698 0 vs 2, p<0.001 0 vs 3, p<0.001 1 vs 2, p<0.001 1 vs 3, p<0.001 2 vs 3, p=0.008
Age groups						
18-30 (n=43)	0 (0%)	2 (4.7%)	11 (25.6%)	30 (69.8%)		
31-40 (n=47)	4 (8.5%)	13 (27.7%)	18 (38.3%)	12 (25.5%)		
41-50 (n=46)	10 (21.7%)	22 (47.8%)	8 (17.4%)	6 (13%)		
Sex						
Female (n=66)	3 (4.5%)	15 (22.7%)	17 (25.8%)	31 (47%)	0.018[b]	
Male (n=70)	11 (15.7%)	22 (31.4%)	20 (28.6%)	17 (24.3%)		
Smoking						
Present (n=76)	9 (11.8%)	23 (30.3%)	23 (30.3%)	21 (27.6%)	0.217[b]	
Absent (n=60)	5 (8.3%)	14 (23.3%)	14 (23.3%)	27 (45%)		
Cumulative smoking (pack-years)	13.5 (23)	7.5 (20)	7 (20)	0 (2)	0.004[c]	0 vs 1, p=0.554 0 vs 2, p=0.548 0 vs 3, p=0.007 1 vs 2, p=0.994 1 vs 3, p=0.005 2 vs 3, p=0.005

Continuous data were presented with mean±SD or median (interquartile range), and categorical data were expressed as frequency (percentage). Differences between groups were analysed using: [a] One-way ANOVA (post-hoc comparison was performed using Tukey's HSD test), [b] Pearson chi-square test, [c] Kruskal-Wallis test (post-hoc comparisons were performed using Dunn's test with Bonferroni correction).

median exposure was significantly lower in patients with a thymic score of 3 compared to those with scores of 0, 1, and 2 (0 vs 13.5, 7.5, and 7 pack-years, respectively; p<0.05) (Table 1).

The transverse diameter of the thymus was significantly larger in the score 3 group (p=0.027). However, AP diameter, length, and thickness of right and left lobes, and overall thymus area did not show any significant difference between thymic score groups (all p>0.05). A statistically significant increase in thymic CT attenuation (HU) was observed with increasing thymic scores (p<0.001). Also, the thymic soft tissue area and soft tissue-to-overall thymus area ratio increased significantly from score 1 to score 3 (p<0.001) (Table 2).

No significant differences were found based on smoking presence for either the thymic score or all other CT-derived parameters (all p>0.05). However, transverse and AP diameters, lobe lengths and thicknesses, and overall thymus area were significantly different between females and males (all p<0.05). The thymic score, density, and soft tissue-to-overall thymus area ratio were significantly lower in male patients (p<0.05) (Table 3).

There was a strong negative correlation between age and thymic score (r= -0.609, p<0.001). In contrast, the thymic score exhibited very strong positive correlations with both CT attenuation (r=0.904, p<0.001) and soft tissue area (r=0.834, p<0.001).

Table 2. CT- derived thymus measurements according to thymic score groups.

	Thymic score 1 (n=37)	Thymic score 2 (n=37)	Thymic score 3 (n=48)	p value	Post-hoc comparison
Transverse diameter (mm)	31 (9)	28 (13)	34 (9)	0.027[b]	1 vs 2, p=0.832 1 vs 3, p=0.017 2 vs 3, p=0.031
Anterior-posterior diameter (mm)	17 (8)	18 (8)	17.5 (6)	0.864[b]	
Right lobe length (mm)	26±7	26±8.5	26±6	0.998[a]	
Right lobe thickness (mm)	10 (4.5)	9 (5)	9 (3.8)	0.475[b]	
Left lobe length (mm)	29.3±7	29.5±7	32.5±7.3	0.064[a]	
Left lobe thickness (mm)	11 (3.8)	11 (4)	11 (2)	0.511[b]	
CT attenuation (HU)	-82.2±12	-36.5±20	14±28.4	<0.001[c]	1 vs 2, p<0.001 1 vs 3, p<0.001 2 vs 3, p<0.001
Overall thymus area (cm^2)	3.5 (2.3)	3.1 (2.42)	3.5 (1.53)	0.706[b]	
Soft tissue area (cm^2)	0.2 (0.22)	1.3 (1.15)	2.96 (1.3)	<0.001[b]	1 vs 2, p<0.001 1 vs 3, p<0.001 2 vs 3, p<0.001
Soft tissue/ overall thymus area ratio	0.07 (0.1)	0.4 (0.32)	0.9 (0.35)	<0.001[b]	1 vs 2, p<0.001 1 vs 3, p<0.001 2 vs 3, p<0.001

Continuous data were presented with mean±SD or median (interquartile range). Quantitative assessments were not applicable for patients with a thymic score of 0. Differences between groups were analysed using: [a] One-way ANOVA, [b] Kruskal-Wallis test. Post-hoc comparisons were performed using Dunn's test with Bonferroni correction, [c] Welch ANOVA, Tamhane was used as a post-hoc test.

Age showed a weak positive correlation with cumulative smoking amount (r=0.237, p=0.007), a strong negative correlation with thymic CT attenuation (r=-0.638, p<0.001), and a moderate negative correlation with thymic soft tissue area (r=-0.568, p<0.001). Cumulative smoking amount had weak negative correlations with thymic score (r=-0.297, p=0.001), thymic CT attenuation (r=-0.287, p=0.002), and soft tissue area (r=-0.219, p=0.019).

Thymic scores were dichotomized into two groups based on the degree of fatty replacement: advanced fatty replacement (>50%, scores 0- 1), and mild fatty replacement (≤50%, scores 2-3). Logistic regression model to predict the dichotomized thymic score groups was statistically significant (X^2 (3)= 62.782, p<0.001, NagelkerkeR^2=0.504). Age (OR=1.23, 95%CI:1.14-1.32, p<0.001) and male sex (OR=4.52, 95%CI: 1.69-12.08, p=0.003) were independent factors related to advanced fatty replacement. Each one-year increase in age was associated with a 1.23-fold and male sex was associated with a 4.52-fold higher odds of advanced fatty replacement. Smoking status had no statistically significant effect (p=0.305).

Discussion

The main findings of our study show that thymic involution is primarily driven by age and sex. Male patients had significantly lower thymic scores and CT attenuation, corresponding to increased fatty replacement. While lower thymic scores were strongly correlated with lower CT attenuation and soft tissue area, most size-related parameters remained similar across score groups, suggesting that fatty replacement rather than true size reduction predominates during thymic involution. Although the cumulative smoking amount differed significantly between thymic score groups and showed negative correlations with thymic score, soft tissue area, and density, logistic regression analysis

Table 3. CT- derived thymus measurements according to sex.

	Female (n=63)	Male (n=59)	p value
Age	37.5 (16)	34.5 (14)	0.882
Cumulative smoking (pack-year)	1 (10)	6.5 (20)	0.107
Thymic score	2 (2)	2 (1)	0.002
Transverse diameter (mm)	30 (9)	33 (9)	0.016
Anterior-posterior diameter (mm)	16 (5)	21 (6)	<0.001
Right lobe length (mm)	23 (9)	28 (8)	<0.001
Right lobe thickness (mm)	8 (3)	10 (3)	0.002
Left lobe length (mm)	29 (10)	32 (9)	0.019
Left lobe thickness (mm)	10 (3)	12 (3)	0.004
CT attenuation (HU)	-12 (74)	-42 (66)	0.003
Overall thymus area (cm^2)	2.8 (1.6)	4 (1.8)	<0.001
Soft tissue area (cm^2)	1.8 (2.6)	1.3 (2.3)	0.094
Soft tissue/ overall thymus area ratio	0.63 (0.78)	0.32 (0.57)	<0.001

Continuous data were presented with median (interquartile range). Quantitative assessments were not applicable for patients with a thymic score of 0. Differences between groups were analysed using the Mann-Whitney U test.

demonstrated that smoking was not an independent predictor of advanced fatty replacement.

Age was significantly different between thymic score groups, and exhibited negative correlations with thymic score, CT attenuation, and thymic soft tissue area, consistent with the literature [3-6]. Naik et al. reported the mean age of complete fatty replacement as 45.4 ± 12.9 years [6]. Simanovsky et al. reported that 71% of participants with a mean age of 52.6 (14-78 years) demonstrated complete fatty replacement. In addition, no solid tissue component was observed after 54 years of age [5]. In our study, among individuals between 41 and 50 years of age, the incidence of thymic scores 0 and 1 was 21.7% and 47.8%, respectively. On the other hand, in the study of Araki et al., which included patients in a wide age range (34-92 years), 9% of participants older than 60 years exhibited a thymic soft tissue component [3]. Moreover, in another study of individuals between 60-69 years of age, residual thymic tissue incidence was reported as 34.8% [7]. These findings reflect the variability of the thymic involution process within the population.

In our study, the transverse diameter of the thymus was significantly larger in patients with a thymic score of 3 (p=0.027), with significantly higher CT attenuation, soft tissue area, and soft tissue-to-overall thymus area ratio (p<0.001). In contrast, AP diameters, lengths, and thicknesses of both thymic lobes and total thymic area were not significantly different between thymic score groups. Similarly, Araki et al. found no significant correlation between thymic size and age, apart from AP diameter and right lobe length in females [3]. This may reflect that thymic size does not substantially decrease, but rather undergoes fatty replacement [11,12]. Additionally, in our study thymic score exhibited very strong positive correlations with both CT attenuation (r=0.904, p<0.001) and soft tissue area (r=0.834, p<0.001), with lower thymic scores corresponding to lower CT attenuation values. Therefore, it may be speculated that visual thymic scoring better reflects the degree of involution.

Male patients had significantly lower thymic scores, CT attenuation, and soft tissue-to-overall thymus area ratio compared to females, despite having significantly larger thymic size in the present study. Similarly, Ackman et al. reported that in adults aged 20-30 years, fatty involution occurs earlier and more rapidly in male patients, although the AP diameter was significantly larger [4]. In Sandstedt et al.'s study, the majority of patients with a thymic score of 0 (64%) were male; this ratio was higher in our study (78.6%)[10]. Also, 64.6%

of patients with a thymic score of 3 were female in our cohort. By comparison, this ratio was reported as 85% by Sandstedt et al. and 86% by Araki et al. [3,10]. Our logistic regression analysis revealed that males were 4.5 times more likely to exhibit advanced fatty replacement compared to females (95%CI: 1.69-12.08, p=0.003). This difference may be explained by the fact that sex steroids play an important role in thymic involution, in which androgens may have a greater influence [13]. On the other hand, females also experience a temporary thymic involution during pregnancy due to increased levels of sex steroids, primarily progesterone. This is thought to be an adaptive immune response to prevent fetal rejection and maintain normal pregnancy [14,15].

Thymic involution leads to an impaired immune function and a decrease in naive T-cell production [8,9]. Sex-based differences in immunity have been reported to be more marked in the elderly, with males exhibiting a greater decrease in naive T cell activity compared to females [16]. It may be speculated that the more pronounced immunological decline in males may be related to the greater degree of thymic involution. Similarly, Sandstedt et al. found that age, male sex, and obesity were associated with complete fatty replacement of the thymus. Furthermore, low CT attenuation values were related to lower thymic activity and depletion of naive T-cells. Therefore, they stated that CT may be valuable in assessing immunological aging [10].

Smoking's effect on thymic involution may be associated with the immunosuppressive effects of tobacco exposure and the systemic oxidative stress it induces. Sandhar et al. reported that thymic output was influenced by sex and smoking, based on their analyses of blood T cell populations and the presence of thymic tissue in mediastinal fat tissue using flow cytometry in patients undergoing cardiothoracic surgery [17]. In the present study, although the majority of patients with a thymic score of 0 (64.3%) had a smoking history, smoking status was not significantly different between thymic score groups and was not an independent predictor of thymic involution. On the contrary, Araki et al. observed a significant association between smoking and advanced fatty replacement. They reported that 64% of patients with a score of 3 were non-smokers, whereas patients with a score of 0 were more commonly smokers with high pack-year exposure (49%) [3]. Likewise, we observed that cumulative smoking amount was negatively correlated with thymic score, thymic CT attenuation, and soft tissue area (p<0.05). The median smoking exposure was highest in patients with a thymic score of 0 and significantly lower in those with a score of 3 (p<0.05). Despite this, the lack of statistical significance of smoking status may have resulted from our small sample size.

Our study has several limitations. First, this was a single-center, retrospective study. Second, the sample size was small. Excluding patients with systemic diseases and those with missing smoking data reduced our sample size, since our institution is a tertiary care referral center with a large population with oncologic diseases, and smoking status was generally not documented in medical records. Third, the cumulative smoking amount may be misleading since it depends on the patient's own statement. Fourth, the study population was limited to 18-50 years of age to monitor the active thymic involution process; however, including a wider age range might yield different results. Fifth, contrast-enhanced CT may have influenced attenuation measurements. Lastly, data regarding menopause and hormonal status in female patients was not available due to the retrospective nature, which could have provided further insights into the sex-based differences of thymic involution. Similarly, information on patients' weight and body mass index was unavailable, and the potential impact of obesity could not be evaluated.

In conclusion, this study highlights the heterogeneity of the thymic involution process, which is mainly affected by age and sex. Although cumulative smoking exposure was related to lower thymic score, smoking was not found to be an independent predictor of thymic involution. Quantitative soft tissue segmentation and density analysis demonstrated that thymic involution causes progressive functional thymic soft tissue loss due to fatty replacement rather than a true size reduction. The awareness of this variability may help clinicians and radiologists better understand the "normal" spectrum of the thymus and distinguish normal physiological involution from hyperplasia. Further prospective longitudinal studies including healthy individuals from general population are needed to better elucidate the dynamics of thymic involution.

Author contributions

Conception and design: S.A.D., G.D.; Data acquisition: S.A.D., İ.T., G.D.; Data analysis: S.A.D., J.K.; Data interpretation: S.A.D., İ.T., J.K., G.D.; Drafting of the

manuscript: S.A.D.; Critical revision of the manuscript: S.A.D., İ.T., J.K., G.D. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Hacettepe University, Faculty of Medicine, Ethics Committee for Health Sciences Research (Date: November 25, 2025, Decision/Protocol No: 2025/22-36). Written informed consent was waived due to the retrospective nature of the study.

Data availability statement

The data supporting the findings of this study are not publicly available due to ethical and privacy restrictions, as they contain sensitive patient information and are subject to institutional regulations.

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

References

[1] Suster S, Rosai J. Histology of the normal thymus. Am J Surg Pathol 1990;14(3):284-303. **[Crossref]**

[2] Nishino M, Ashiku SK, Kocher ON, Thurer RL, Boiselle PM, Hatabu H. The thymus: a comprehensive review. Radiographics 2006;26(2):335-48. **[Crossref]**

[3] Araki T, Nishino M, Gao W, et al. Normal thymus in adults: appearance on CT and associations with age, sex, BMI and smoking. Eur Radiol 2016;26(1):15-24. **[Crossref]**

[4] Ackman JB, Kovacina B, Carter BW, et al. Sex difference in normal thymic appearance in adults 20-30 years of age. Radiology 2013;268(1):245-53. **[Crossref]**

[5] Simanovsky N, Hiller N, Loubashevsky N, Rozovsky K. Normal CT characteristics of the thymus in adults. Eur J Radiol 2012;81(11):3581-6. **[Crossref]**

[6] Naik S, Shri A, Sidhu S, et al. Multidetector computed tomography evaluation of normal thymus and variations with age. J Minim Access Surg 2025;21(2):101-6. **[Crossref]**

[7] Drabkin MJ, Meyer JI, Kanth N, et al. Age-stratified patterns of thymic involution on multidetector CT. J Thorac Imaging 2018;33(6):409-16. **[Crossref]**

[8] Li L, Xu F, Han Y, Zeng J, Du S, Wang C. Thymic microenvironment's impact on immunosenescence. Immunol Res 2024;72(5):1161-73. **[Crossref]**

[9] Liang Z, Dong X, Zhang Z, Zhang Q, Zhao Y. Age-related thymic involution: mechanisms and functional impact. Aging Cell 2022;21(8):e13671. **[Crossref]**

[10] Sandstedt M, Chung RWS, Skoglund C, et al. Complete fatty degeneration of thymus associates with male sex, obesity and loss of circulating naïve CD8+ T cells in a Swedish middle-aged population. Immun Ageing 2023;20(1):45. **[Crossref]**

[11] Moore AV, Korobkin M, Olanow W, et al. Age-related changes in the thymus gland: CT-pathologic correlation. AJR Am J Roentgenol 1983;141(2):241-6. **[Crossref]**

[12] Nasseri F, Eftekhari F. Clinical and radiologic review of the normal and abnormal thymus: pearls and pitfalls. Radiographics 2010;30(2):413-28. **[Crossref]**

[13] Gui J, Mustachio LM, Su DM, Craig RW. Thymus size and age-related thymic involution: early programming, sexual dimorphism, progenitors and stroma. Aging Dis 2012;3(3):280-90.

[14] Tibbetts TA, DeMayo F, Rich S, Conneely OM, O'Malley BW. Progesterone receptors in the thymus are required for thymic involution during pregnancy and for normal fertility. Proc Natl Acad Sci U S A 1999;96(21):12021-6. **[Crossref]**

[15] Erlandsson MC, Ohlsson C, Gustafsson JA, Carlsten H. Role of oestrogen receptors alpha and beta in immune organ development and in oestrogen-mediated effects on thymus. Immunology 2001;103(1):17-25. **[Crossref]**

[16] Márquez EJ, Chung CH, Marches R, et al. Sexual-dimorphism in human immune system aging. Nat Commun 2020;11(1):751. **[Crossref]**

[17] Sandhar B, Vyas V, Harding D, et al. Heterogeneity of thymic output in the elderly and its association with sex and smoking. JCI Insight 2025;10(15):e189008. **[Crossref]**

Acta Medica 2026; 57(3): 220-226 DOI: 10.32552/actamedica.2026.1305

acta medica RESEARCH ARTICLE

Recognizable clinical features of 1p36 deletion syndrome: Insights from a single-center cohort

Gizem Ürel Demir[1], Serap Ketenci İşlek[1], Gülen Eda Utine[1], Pelin Özlem Şimşek Kiper[1]

[1]Division of Pediatric Genetics, Department of Pediatrics, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Abstract

Objective: 1p36 deletion syndrome is one of the most common contiguous gene deletion syndromes and is primarily characterized by a typical facial gestalt, as well as neurological problems and multisystemic involvement. This study presents the clinical findings of 17 patients with 1p36 deletion syndrome followed at a single center.

Methods: Data from patients diagnosed with 1p36 deletion syndrome between 2010 and 2026 were retrospectively reviewed. Demographic characteristics, growth parameters, dysmorphic features, and neurological, cardiac, endocrine, visual, and auditory findings were analyzed, and the diagnostic methods used were evaluated.

Results: A total of 17 patients (13 females and 4 males) were included in the study. Four patients were diagnosed using microarray analysis, and 13 using fluorescence in situ hybridization (FISH). Neurodevelopmental delay was observed in all patients. Typical facial gestalt was observed in the patients, with the most common dysmorphic facial features being straight eyebrows, deep-set eyes, and ear abnormalities. Central nervous system malformations were common, and epilepsy was observed in 70% of patients. Cardiac findings, ocular problems, hearing impairments, and hypothyroidism were also frequently identified. Three patients died, all due to respiratory failure.

Conclusion: 1p36 deletion syndrome is a clinically recognizable condition, and early diagnosis and referral for appropriate management are important for improving prognosis. Although phenotypes associated with variable deletion sizes and genes located within critical regions have been reported, a multisystem approach is required in this syndrome due to its variable clinical presentation.

Keywords: 1p36 deletion syndrome, dysmorphism, developmental delay, epilepsy, fluorescence in situ hybridization, microarray

Introduction

1p36 deletion syndrome is among the most common contiguous gene deletion syndromes with an estimated prevalence of 1 in 5,000 live births [1]. The disorder was first described in two unrelated children with an unbalanced t(1;15) translocation causing loss of chromosomal material from 1p and 15q [2]. Subsequently, both isolated 1p36 deletions and 1p36 monosomy resulting from unbalanced translocations with gain of material from another chromosome were described in the literature. Comparison of patients with isolated deletions involving the short arm of chromosome 1 helped establish the characteristic phenotypic gestalt of the syndrome [3]. More recent data indicate that the majority of cases arise de novo, while only a small proportion are associated with parental balanced rearrangements [4]. Furthermore, recurrent interstitial deletions of maternal origin have

Corresponding author: Gizem Ürel Demir • **Email:** gizemurel@gmail.com

Received: April 20, 2026 **Accepted:** July 31, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

been reported in siblings, supporting the possibility of underlying germline mosaicism [4].

Monosomy 1p36 is characterized by a spectrum of neurological abnormalities, developmental delay, intellectual disability, congenital anomalies, short stature, and distinctive facial features [5]. The phenotype is mainly attributed to haploinsufficiency of genes within the deleted region. Although the deleted chromosome has been reported to be more commonly of maternal origin, no clear parent-of-origin effect on the phenotype has been demonstrated [6]. The disorder can often be clinically recognized by its characteristic facial appearance. Common features include deep-set eyes, straight eyebrows, midface hypoplasia, a flat nasal bridge, a long philtrum, a pointed chin, cleft lip and/or palate, and ear abnormalities [6]. However, the clinical presentation varies considerably among affected individuals.

Terminal deletions have been reported in more than half of the patients (52–67%). In the remaining patients, deletions were observed to arise from interstitial deletions (10–29%), complex rearrangements (12%), or derivative chromosomes associated with unbalanced translocations (7–16%) [5,7]. Although most deletions detectable by conventional chromosome analysis are located in terminal chromosomal regions, deletions involving 1p36 may be difficult to identify [6]. Accordingly, molecular cytogenetic techniques such as fluorescence in situ hybridization (FISH) or chromosomal microarray analysis (CMA) are preferred for diagnosis.

In this study, we present the clinical and cytogenetic characteristics of 17 patients diagnosed with 1p36 deletion syndrome followed at a single center. To the best of our knowledge, this study represents the first and a relatively large cohort of patients with 1p36 deletion syndrome reported from Türkiye.

Materials and Methods

Patients diagnosed with 1p36 deletion syndrome at the Department of Pediatric Genetics of Hacettepe University, between 2010 and 2026, based on FISH and/or CMA, were included in the study. Patient data were retrospectively reviewed from hospital and laboratory records. One of the patients included in this cohort has been previously reported [8].

Demographic characteristics of the patients, including sex, age at presentation, age at diagnosis, current age, and parental ages at birth, were collected. Age at death was additionally recorded for deceased patients. The presence of prenatal ultrasonographic abnormalities was assessed. Anthropometric measurements of patients at the time of admission were examined according to age- and sex-adjusted standard deviation scores, and short stature and microcephaly were determined. The frequency of dysmorphic facial features was also assessed. Data were collected on neurological, cardiac, ophthalmological, auditory, and endocrine manifestations. Denominators varied across variables because some clinical records were incomplete and certain examinations or investigations had not been performed or were unavailable for all patients. Percentages were calculated based on the number of patients with available data for each variable.

FISH analysis was performed using the Vysis 1p36 Microdeletion Syndrome Probe Set, targeting the CDC2L1/p58 locus at 1p36, with a control probe at 1q25 (Abbott Molecular, IL, USA), and the CytoCell LPU 020 Monosomy 1p36 Probe, targeting the SKI locus at 1p36.33 with a 1qter control probe (Oxford Gene Technology, UK). CMA was performed using the CytoScan Optima oligonucleotide array (Thermo Fisher Scientific, Waltham, MA, USA) and Agilent ISCA 8x60K CGH Array (Agilent Technologies, Santa Clara, CA, USA).

The study was approved by the Hacettepe University Ethics Committee (SBA 26/270). Written informed consent for the publication of identifiable clinical photographs was obtained from the parents or legal guardians of all patients shown in Figure 1.

Results

A total of 17 patients were included in the study. Four were diagnosed using microarray analysis, while the remaining patients were diagnosed using FISH following a preliminary diagnosis of 1p36 deletion syndrome (Figure 2). Among the cases diagnosed by microarray, three of four had no prior clinical suspicion, whereas one patient with a Robertsonian translocation, 45,XX,rob(13;14)(q10;q10), underwent microarray analysis despite having a preliminary diagnosis. Conventional karyotyping was performed in 14 patients, and a 1p36 deletion was detected in only one of them.

The deletion size detected by microarray analysis ranged from 3685 kb to 15.7 Mb.

The mean age at admission and at diagnosis was 22 and 28 months, respectively. Prenatal CNS anomalies, including ventriculomegaly and corpus callosum hypoplasia, were detected in two patients. Growth impairment and microcephaly were common at presentation. Dysmorphic facial features were present in all patients, most commonly straight eyebrows, deep-set eyes, and ear abnormalities (Figure 1). Extremity abnormalities, including clubfoot, digital anomalies, and developmental hip dysplasia, were observed in eight patients.



Figure 1. Clinical photographs of three patients diagnosed with 1p36 deletion syndrome.

Facial photographs of Patient 1 at 1 year (a) and 4 years (b); Patient 2 at 4 years (c); and Patient 3 at 4 years (d) and 9 years (e), demonstrating characteristic facial features including a prominent forehead, straight eyebrows, and deep-set eyes. Hand photograph of Patient 3 (f) showing a broad hand with wide, deep-set nails.

Developmental delay was present in all patients. CNS abnormalities commonly included corpus callosum abnormalities, ventriculomegaly, and polymicrogyria. Congenital heart defects included atrial septal defect (ASD), ventricular septal defect (VSD), patent ductus arteriosus (PDA), bicuspid aortic valve, and aortic regurgitation, while noncompaction cardiomyopathy was also observed. Ophthalmological abnormalities included myopia, cataract, and optic disc pallor. Hearing loss, hypothyroidism, swallowing dysfunction, and cleft lip and palate were also identified.

Three of the 17 patients died, and three others could not be reached for long-term follow-up. All three children died of respiratory failure, two during infancy and one in the neonatal period. At the latest evaluation, severe neurodevelopmental impairment was evident among patients with available follow-up data. Only 4 of 11 patients (36.4%) achieved independent walking, and all did so at or after 24 months of age, while 9 of 11

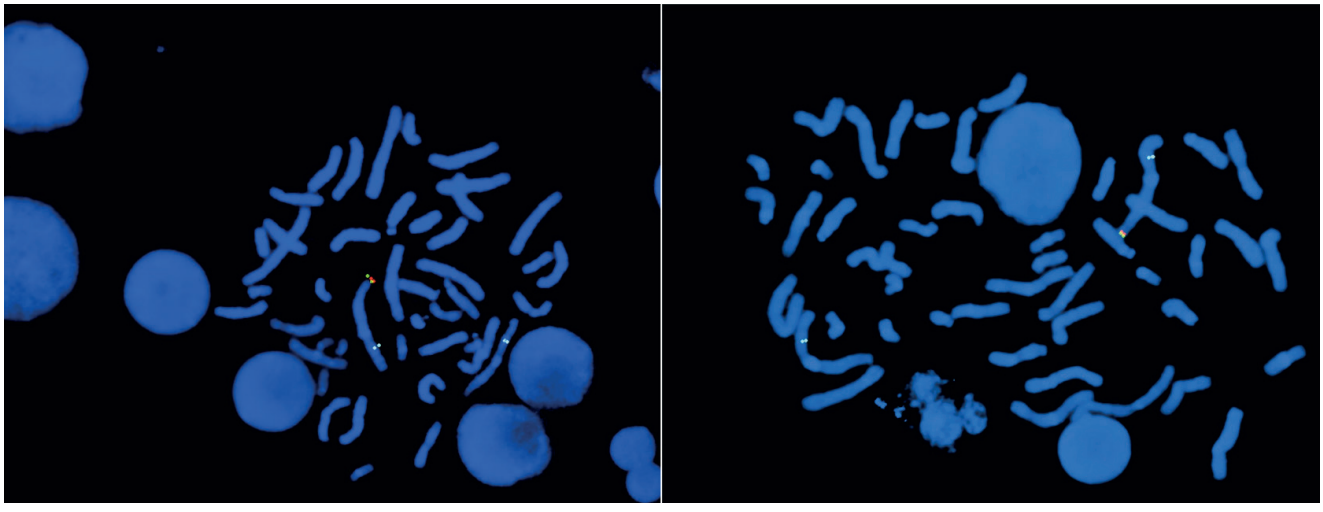


Figure 2. Representative FISH images demonstrating loss of one 1p36 target signal with retention of both control signals, consistent with a heterozygous 1p36 deletion.

Table 1. Summary of demographic characteristics, clinical findings, and diagnostic methods in patients with 1p36 deletion syndrome.

Parameter	Value
Demographic characteristics	
Sex (F/M)	13 (76.5%) / 4 (23.5%)
Age at presentation (months)	Mean 22 (range newborn-52)
Age at diagnosis (months)	Mean 28 (range 6-66)
Current age (years)	Mean 14.3 (range 9.5-18)
Maternal age at delivery (years)	26.3 (range 17-42)
Paternal age at delivery (years)	30.9 (range 25-43)
Diagnostic method	
FISH	13/17 (76.5%)
Chromosomal microarray	4/17 (23.5%)
Prenatal ultrasound abnormalities	2/16 (12.5%)
Anthropometric measurements	
Height SDS	-0.77 ± 1,27
Weight SDS	-1.58 ± 2.41
Head circumference SDS	-1.86 ± 1.37
Short stature	5/16 (31.2%)
Microcephaly	13/17 (76.4%)
Facial dysmorphism	
Large anterior fontanel	2/6 (33.3%)
Brachycephaly	8/12 (66.6%)
Prominent forehead	11/17 (64.7%)
Hypertelorism	3/17 (17.6%)
Straight eyebrows	14/16 (87.5%)
Deep-set eyes	14/17 (82.3%)
Epicanthus	2/17 (11.7%)
Depressed nasal bridge	8/17 (47%)
Abnormal ears	13/16 (81.2%)
Long philtrum	5/16 (31.2%)
Micrognathia	5/17 (29.4%)
Limb abnormalities	8/17 (47%)
Developmental delay	16/16 (100%)
CNS abnormalities	13/16 (81.2%)
Epilepsy	12/17 (70.5%)
Congenital heart defects	5/16 (31.2%)
Cardiomyopathy	2/16 (12.5%)
Ophthalmological abnormalities	3/15 (20%)
Hearing loss	2/14 (14.2%)
Hypothyroidism	4/8 (50%)

CNS: central nervous system, F: female, FISH: fluorescence in situ hybridization, M: male, SDS: standard deviation scores.

patients (81.8%) were unable to perform daily activities without assistance. Also, language development was significantly affected, 4 of 11 patients (36.4%) remained non-verbal, while the others had a limited vocabulary (5–10 words) or were able to form only simple two-word sentences. Behavioral problems, particularly irritability, aggression, and autistic features, were detected in 5 of 11 patients (45.5%). Additionally, scoliosis was detected in 3 of 11 patients (27.3%) and obesity and/or hyperphagia in 4 of 11 patients (36.4%), with these findings becoming more pronounced after the age of 10. The demographic and clinical characteristics of the patients are summarized in Table 1.

Discussion

Monosomy 1p36 is recognized as the most prevalent terminal chromosomal deletion in humans, which accounts for 0.5–1.2% of syndromic intellectual disability cases [7]. The phenotypic differences observed between affected individuals were initially thought to occur due to the parent-of-origin effect, but a study investigating this through DNA polymorphism analysis showed that there was no such effect [3]. The same study suggested that the size of 1p36 deletions varies significantly among patients and that these phenotypic differences may be due to deletions of varying sizes that result in the loss of different genes [3]. Moreover, differences in craniofacial features were not related solely to the size of the chromosomal deletion [3].

The deletion breakpoints and size show significant variability among patients and have been shown not to always correlate with the severity of the phenotype [9]. In a recent study, cases stratified into two groups as proximal or distal based on deletion location showed that cardiac abnormalities, CNS malformations, and epilepsy were more frequent in distal deletions, whereas microcephaly was more common in proximal deletions [5]. Additionally, deletions larger than 10 Mb have been shown to be associated with more severe neurodevelopmental outcomes [5]. In our study, due to the small number of patients in whom breakpoints were clearly identified by microarray analysis, it was not possible to establish such a correlation. However, microarray analysis in the two deceased patients showed deletions of different sizes, with a 15.7 Mb

deletion in one and a 6922 kb deletion in the other. Conventional karyotyping detected the 1p36 deletion in only 1 of 14 patients tested in our cohort, indicating its limited diagnostic sensitivity for this syndrome. Therefore, when 1p36 deletion syndrome is clinically suspected, targeted FISH or CMA should be pursued even if the conventional karyotype is normal.

Previous studies suggest a female predominance; similarly, in our cohort, the female:male ratio was 13:4. More than 75% of patients exhibited characteristic dysmorphic facial features, including straight eyebrows, deep-set eyes, midface hypoplasia, a flat nasal bridge, posteriorly rotated, low-set, and asymmetric ears, a long philtrum, and a pointed chin [7,10]. Additionally, microcephaly and a large or late-closing anterior fontanel are frequently observed. In our cohort, a characteristic facial gestalt was observed, particularly marked by straight eyebrows and deep-set eyes, similar to those described in the literature. In 14 out of 17 patients, we had a preliminary diagnosis, indicating that characteristic craniofacial features can effectively guide the diagnosis. Of the three patients for whom facial features did not initially lead to a diagnosis, one was evaluated as a newborn and the other two were around 2 years old. It was noted that dysmorphic features became more pronounced in patients over time. Haploinsufficiency of MMP23B (OMIM* 603321) has been proposed to contribute to a large or late-closing anterior fontanel, while SKI (OMIM* 164780) has been suggested as a candidate gene for the cleft lip/palate phenotype based on Ski-null mouse models [11,12].

Developmental delay and/or intellectual disability are observed in the vast majority of patients with 1p36 deletion syndrome, and haploinsufficiency of genes located within this locus, such as GNB1 (OMIM* 139380), CHD5 (OMIM* 610771), KCNAB2 (OMIM* 601142), SKI, GABRD (OMIM* 137163), MTOR (OMIM* 601231), RERE (OMIM* 605226), and CDC42 (OMIM* 116952), is thought to contribute to the neurodevelopmental phenotype [1,5]. All of our patients showed significant neurodevelopmental impairment based on clinical observations; however, formal cognitive assessment was not performed.

Epilepsy has been reported in 44–70% of cases, with both generalized and focal seizure types described, although infantile spasms are the most frequent [5,13,14]. Genotype–phenotype studies have shown that epilepsy is more frequent in distal deletions and that deletions involving GABRD and KCNAB2 are associated with the epilepsy phenotype [5]. Epilepsy was a prominent neurological feature in our cohort, observed in 70% of the patients.

CNS abnormalities, most notably ventriculomegaly and corpus callosum abnormalities, as well as cortical atrophy, periventricular nodular heterotopia, polymicrogyria, and white matter abnormalities, have been described [5,14-16]. In our cohort, consistent with the literature, ventriculomegaly and corpus callosum abnormalities were the most frequently observed findings. These anomalies may also be detectable during the prenatal period and should be kept in mind during evaluation.

Cardiac pathologies have been identified as the leading cause of mortality in the literature [5]. In the largest recent cohort, cardiovascular malformations or cardiomyopathy were reported in 47% of patients, including structural heart defects in 32% and cardiomyopathy in 21%, whereas the earlier Battaglia series reported congenital heart defects in 71% and cardiomyopathy in 27% [5,14]. While PRDM16 (OMIM*605557) and SKI have been proposed as the main genes implicated in cardiomyopathy, the relevant locus contains numerous genes, such as ECE1 (OMIM* 600423) and RERE, that are thought to contribute to the pathogenesis of cardiac defects [1,5]. HSPB7 (OMIM* 610692) has also been suggested as a potential candidate gene for cardiomyopathy based on population-level genetic association studies [17,18]. The frequency of congenital heart defects in our cohort was comparable to that reported in the largest recent cohort, whereas mortality was due to respiratory failure rather than cardiac causes. Among the patients who died, one had a large ASD, another had a bicuspid aortic valve with aortic regurgitation, and the third had no congenital heart defect; none showed cardiomyopathy.

Hypothyroidism has been reported, and screening for hypothyroidism is recommended at birth, at 6 months, and annually thereafter [10]. Hypothyroidism was observed in 4/8 (50%) of our patients, and since data for the remaining patients were unavailable, a precise frequency could not be determined; therefore, selection bias may exist.

Prader-Willi syndrome is included in the differential diagnosis, as some patients develop obesity after early feeding difficulties [19,20]. Among our patients, two were treated for swallowing dysfunction, and four developed obesity and/or hyperphagia during follow-up.

The main limitation of our study is that microarray analysis was not performed in all patients to precisely determine the breakpoints, which precluded the establishment of a genotype/phenotype correlation. Furthermore, due to the lack of parental studies, it could not be determined whether the deletion was de novo. Missing data for several clinical variables, particularly endocrine, ophthalmological, auditory, and CNS findings, led to varying denominators and may have affected the reported frequencies. Early diagnosis, along with timely screening and management of associated comorbidities, particularly neurological ones, is crucial for improving patient outcomes. Recognition of patients with a typical facial gestalt may facilitate earlier diagnosis and prevent unnecessary diagnostic testing. The possibility of balanced parental rearrangement should be considered, especially in cases with additional chromosomal copy number variations. Furthermore, germline mosaicism should be taken into account when providing genetic counseling to families regarding future pregnancies.

Author contributions

Conception and design: G.Ü.D., G.E.U., P.Ö.Ş.K.; Data acquisition: S.K.İ.; Data analysis: G.Ü.D., S.K.İ., P.Ö.Ş.K.; Data interpretation: G.Ü.D., S.K.İ., P.Ö.Ş.K.; Drafting of the manuscript: G.Ü.D.; Critical revision of the manuscript: G.Ü.D., S.K.İ., G.E.U., P.Ö.Ş.K. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Hacettepe University Ethics Committee (Date: 23.03.2026, Decision/Protocol No: SBA 26/270). Informed consent was obtained from all participants involved in this study.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

References

[1] Jordan VK, Zaveri HP, Scott DA. 1p36 deletion syndrome: an update. Appl Clin Genet 2015;8:189-200. **[Crossref]**

[2] Hain D, Leversha M, Campbell N, Daniel A, Barr PA, Rogers JG. The ascertainment and implications of an unbalanced translocation in the neonate. Familial 1:15 translocation. Aust Paediatr J 1980;16(3):196-200. **[Crossref]**

[3] Shapira SK, McCaskill C, Northrup H, et al. Chromosome 1p36 deletions: the clinical phenotype and molecular characterization of a common newly delineated syndrome. Am J Hum Genet 1997;61(3):642-50. **[Crossref]**

[4] Gajecka M, Saitta SC, Gentles AJ, et al. Recurrent interstitial 1p36 deletions: Evidence for germline mosaicism and complex rearrangement breakpoints. Am J Med Genet A 2010;152A(12):3074-83. **[Crossref]**

[5] Jacquin C, Landais E, Poirsier C, et al. 1p36 deletion syndrome: review and mapping with further characterization of the phenotype, a new cohort of 86 patients. Am J Med Genet A 2023;191(2):445-58. **[Crossref]**

[6] Heilstedt HA, Ballif BC, Howard LA, Kashork CD, Shaffer LG. Population data suggest that deletions of 1p36 are a relatively common chromosome abnormality. Clin Genet 2003;64(4):310-6. **[Crossref]**

[7] Guterman S, Beneteau C, Redon S, et al. Prenatal findings in 1p36 deletion syndrome: new cases and a literature review. Prenat Diagn 2019;39(10):871-82. **[Crossref]**

[8] Akkuş PZ, Şahin Y, Utine E, Boduroğlu K. 1p36 microdeletion syndrome: a case report. Acta Medica. 2014;45(1):26-8.

[9] Rocha CF, Vasques RB, Santos SR, Paiva CLA. Mini-review: monosomy 1p36 syndrome: reviewing the correlation between deletion sizes and phenotypes. Genet Mol Res 2016;15(1):10.4238/gmr.15017942. **[Crossref]**

[10] Heilstedt HA, Ballif BC, Howard LA, et al. Physical map of 1p36, placement of breakpoints in monosomy 1p36, and clinical characterization of the syndrome. Am J Hum Genet 2003;72(5):1200-12. **[Crossref]**

[11] Colmenares C, Heilstedt HA, Shaffer LG, et al. Loss of the SKI proto-oncogene in individuals affected with 1p36 deletion syndrome is predicted by strain-dependent defects in Ski-/- mice. Nat Genet 2002;30(1):106-9. **[Crossref]**

[12] Gajecka M, Yu W, Ballif BC, et al. Delineation of mechanisms and regions of dosage imbalance in complex rearrangements of 1p36 leads to a putative gene for regulation of cranial suture closure. Eur J Hum Genet 2005;13(2):139-49. **[Crossref]**

[13] Bahi-Buisson N, Guttierrez-Delicado E, Soufflet C, et al. Spectrum of epilepsy in terminal 1p36 deletion syndrome. Epilepsia 2008;49(3):509-15. **[Crossref]**

[14] Battaglia A, Hoyme HE, Dallapiccola B, et al. Further delineation of deletion 1p36 syndrome in 60 patients: a recognizable phenotype and common cause of developmental delay and mental retardation. Pediatrics 2008;121(2):404-10. **[Crossref]**

[15] Descartes M, Mikhail FM, Franklin JC, McGrath TM, Bebin M. Monosomy1p36.3 and trisomy 19p13.3 in a child with periventricular nodular heterotopia. Pediatr Neurol 2011;45(4):274-8. **[Crossref]**

[16] Dobyns WB, Mirzaa G, Christian SL, et al. Consistent chromosome abnormalities identify novel polymicrogyria loci in 1p36.3, 2p16.1-p23.1, 4q21.21-q22.1, 6q26-q27, and 21q2. Am J Med Genet A 2008;146A(13):1637-54. **[Crossref]**

[17] Matkovich SJ, Van Booven DJ, Hindes A, et al. Cardiac signaling genes exhibit unexpected sequence diversity in sporadic cardiomyopathy, revealing HSPB7 polymorphisms associated with disease. J Clin Invest 2010;120(1):280-9. **[Crossref]**

[18] Stark K, Esslinger UB, Reinhard W, et al. Genetic association study identifies HSPB7 as a risk gene for idiopathic dilated cardiomyopathy. PLoS Genet 2010;6(10):e1001167. **[Crossref]**

[19] Tsuyusaki Y, Yoshihashi H, Furuya N, et al. 1p36 deletion syndrome associated with Prader-Willi-like phenotype. Pediatr Int 2010;52(4):547-50. **[Crossref]**

[20] D'Angelo CS, Da Paz JA, Kim CA, et al. Prader-Willi-like phenotype: investigation of 1p36 deletion in 41 patients with delayed psychomotor development, hypotonia, obesity and/or hyperphagia, learning disabilities and behavioral problems. Eur J Med Genet 2006;49(6):451-60. **[Crossref]**

Acta Medica 2026; 57(3): 227-237 DOI: 10.32552/actamedica.2026.1306

acta medica RESEARCH ARTICLE

Determinants of functional recovery after tibial plateau fractures: The role of fracture severity, fixation, and complexity

Gökhan Ayık[1], Berat Yüksel[1], Orhan Mete Karademir[2], Ulaş Can Kolaç[1]

[1]Department of Orthopedics and Traumatology, Hacettepe University, Ankara, Türkiye
[2]Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Abstract

Objective: Tibial plateau fractures are complex intra-articular injuries with variable functional outcomes. While fracture classifications guide management, their ability to predict recovery remains unclear. This study evaluated the impact of fracture severity, treatment variables, and fracture complexity on patient-reported and objective outcomes over time.

Methods: A retrospective cohort of 75 patients who underwent surgical treatment for tibial plateau fractures was analyzed. Functional outcomes were assessed using the Knee Injury and Osteoarthritis Outcome Score (KOOS) and knee flexion at 1 month, 3 months, 6 months, 1 year, and 2 years postoperatively. Fractures were classified using Schatzker and AO/OTA systems. Fixation method, column and segment involvement, and patient factors were recorded. Longitudinal changes were analyzed using generalized estimating equations, and predictors of early improvement were evaluated using robust regression.

Results: KOOS and knee flexion improved significantly over time ($p<0.001$). Mean KOOS increased from 23.15±5.87 at 1 month to 84.19±8.74 at 2 years, while flexion improved from 74.12±13.91° to 120.91±17.31°. Schatzker and AO/OTA classifications were associated with KOOS ($p<0.001$) but not flexion ($p=0.160$). Definitive internal fixation method was not significantly associated with KOOS outcomes ($p=0.817$). In robust regression, male sex ($\beta=4.146$, $p=0.042$) and posterior segment count ($\beta=2.011$, $p=0.022$) were positively associated with KOOS improvement, whereas comorbidity ($\beta=-7.245$, $p=0.012$) and provisional external fixation before definitive surgery ($\beta=-8.683$, $p<0.001$) were associated with less improvement, statistically.

Conclusion: Recovery following tibial plateau fractures is primarily time-dependent, with fracture severity influencing patient perception but not objective motion. No significant associations were identified between functional outcomes and fixation method or the evaluated fracture-complexity parameters in the present cohort.

Keywords: tibial plateau fractures, KOOS, fracture classification, range of motion, outcome assessment

Introduction

Tibial plateau fractures are complex intra-articular injuries that can significantly affect knee function and long-term joint health [1,2]. Restoration of articular congruity, mechanical alignment, and joint stability are considered essential for optimal recovery [1]. Despite advances in surgical techniques and fixation methods, functional outcomes remain variable, and the factors influencing recovery are not fully understood [2,3].

Corresponding author: Gökhan Ayık • **Email:** drgayik@gmail.com

Received: April 23, 2026 **Accepted:** July 23, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Fracture classification systems such as Schatzker and AO/OTA are widely used to describe injury patterns and guide treatment strategies [1,2]. In general, more severe fractures are thought to be associated with worse clinical outcomes [4,5]. At the same time, considerable emphasis has been placed on surgical technique, with the assumption that fixation method may influence postoperative recovery [3,6]. However, the extent to which fracture severity and fixation strategy independently affect functional outcomes remains unclear [1,2,4].

More recently, concepts such as the 3-column model and ten-segment classification have been introduced to better characterize fracture complexity and guide surgical planning [7-9]. Although these systems provide a more detailed anatomical description, their relationship with clinical outcomes has not been clearly established [8,9]. In addition, it is not well defined whether different outcome domains, such as patient-reported measures and objective range of motion, are influenced by the same factors [9,10].

Therefore, the aim of this study was to evaluate the impact of fracture severity, fixation method, and fracture complexity on functional outcomes following tibial plateau fractures. We hypothesized that fracture severity would negatively affect both patient-reported and objective outcomes, fixation method would have a significant impact on functional recovery.

Methods

The research was conducted in accordance with the principles of the Declaration of Helsinki. Ethical approval was obtained from the institutional review board (IRB No: SBA 25/682).

This retrospective cohort study included consecutive patients who underwent surgical treatment for tibial plateau fractures at a single tertiary referral center between 2014-2023. Patients were eligible for inclusion if they were aged ≥18 years, had a tibial plateau fracture treated surgically, and had available clinical and radiological follow-up data. Patients with pathological fractures, prior surgery on the affected knee, or incomplete medical records were excluded, Figure 1. All 75 included patients had available KOOS and knee flexion measurements at the predefined follow-up assessments.

Data collection

Demographic and clinical data were obtained from institutional records, including age at the time of surgery, sex, affected side, presence of comorbidities, trauma mechanism. Surgical variables included fixation method, use of bone graft (autograft, allograft, or none), application of external fixation, reoperation status. In the fixation classification, “plate plus screw(s)” indicated plate fixation combined with one or more additional independent lag screws placed outside the plate.

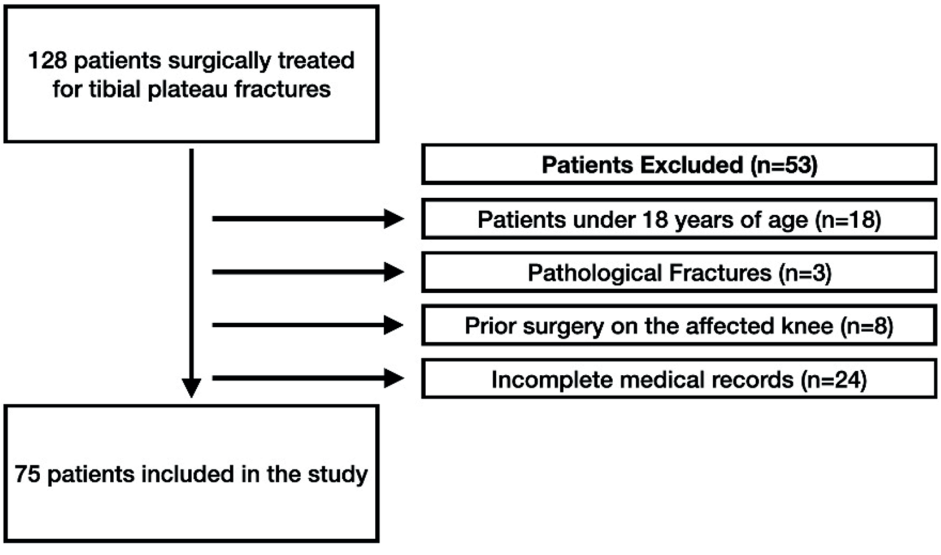


Figure 1. Inclusion and exclusion diagram.

Fracture classification

Fractures were classified using preoperative imaging, including plain radiographs and computed tomography scans. Fracture patterns were categorized according to the Schatzker and AO/OTA classification systems [7-9]. Fracture complexity was further assessed using the 3-column concept and the ten-segment classification. The number of involved columns, as well as anterior and posterior segment involvement, was recorded for each patient.

Radiological assessment

Radiological measurements were performed using the institutional picture archiving and communication system (PACS). Preoperatively measured parameters included articular depression (mm), step-off (mm), plateau widening (mm), fracture gap (mm), varus/valgus alignment (degrees), and posterior tibial slope (degrees). All radiological measurements were performed by a single observer who was blinded to the clinical outcome data. Each parameter was measured independently on two separate occasions, and the mean of the two measurements was used for analysis.

Outcome measures

Functional outcomes were assessed using the validated Turkish version of the Knee Injury and Osteoarthritis Outcome Score (KOOS) and knee flexion range of motion measured in degrees. KOOS was calculated according to the standard scoring instructions, with higher scores indicating better knee-related outcomes. For the present analysis, an overall KOOS score was derived by calculating the global mean of the five individual validated subscale scores. Evaluations were performed at standardized postoperative time points: 1 month, 3 months, 6 months, 1 year, and 2 years [4,10-12]. Change scores for KOOS and knee flexion were calculated as the difference between the 2-year and 1-month postoperative values (2-year value minus 1-month value).

Statistical analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, Version 23.0 (IBM Corp., Armonk, NY, USA) and R software. Normality of continuous variables was assessed using the Kolmogorov–Smirnov test. Descriptive statistics were presented as mean ± standard deviation for normally distributed continuous variables, median (minimum–maximum) for non-normally distributed continuous variables, and frequency (percentage) for categorical variables. The effects of independent variables on non-normally distributed change scores were analyzed using robust regression analysis with the MASS package in R. To evaluate the effects of fixation method, Schatzker classification, AO/OTA classification, time, 3-column concept, posterior segment count, and anterior segment count on repeated KOOS and knee flexion measurements, generalized estimating equations (GEE) were used. Both main effects and time interactions were assessed in the GEE models. For variables with significant effects, multiple comparisons were performed using Bonferroni correction. For comparisons of parameters within the same time point based on the multi-response structure of the 3-column concept and ten-segment classifications, Bonferroni-adjusted multiple comparison tests were applied. Longitudinal changes over time were analyzed using generalized estimating equations (GEE), which account for within-subject correlations. A p-value <0.05 was considered statistically significant for all analyses.

Results

A total of 75 patients were included in the study, of whom 55 (73%) were male and 20 (27%) were female. The mean age at the time of surgery was 45.20±16.74 years. The most common trauma mechanism was fall (41%), followed by in-vehicle traffic accidents (33%). The right side was affected in 45% of patients. According to fracture classification, Schatzker type 6 (30%) and AO/OTA type 41-C (35%) fractures were the most frequent. Single plate with additional screw(s) was the most commonly used fixation method (31%). Bone graft was not used in 54.6% of cases. External fixation was applied in 19% of patients (Table 1). The complications were arthrofibrosis, infection, nonunion and implant failure.

Both KOOS and knee flexion improved significantly over time. The mean KOOS increased from 23.15±5.87 at 1 month to 84.19±8.74 at 2 years (p<0.001). Similarly, mean knee flexion improved from 74.12±13.91° to 120.91±17.31° over the same period (p<0.001). Overall, improvements in both KOOS and knee flexion were statistically significant (p<0.001) (Table 2 and Table 3).

Table 1. Demographic characteristics and clinical profile of patients.

Variables		Value
Age		45.20 ± 16.74
Sex	Male	55 (73%)
	Female	20 (27%)
Side	Right	34 (45%)
	Left	41(55%)
Comorbidities	Present	11 (15%)
	None	64 (85%)
Trauma Mechanism	Fall	31 (41%)
	In-Vehicle accident	25 (33%)
	Motorcycle accident	12 (16%)
	Pedestrian traffic accident	5 (7%)
	Scooter accident	2 (3%)
Fixation Method	Single plate with additional screw(s)	23 (31%)
	Single Plate	19 (25%)
	Multiple plates with additional screw(s)	18 (24%)
	Screw fixation alone	8 (11%)
	Multiple Plates	7 (9%)
Bone Graft Used	Autograft	14 (19%)
	Allograft	20 (27%)
	None	41 (54%)
External Fixation Before Definitive Surgery		14 (19%)
Reoperation Occurred		6 (8%)
Schatzker Classification	1	13 (17%)
	2	7 (9%)
	3	16 (21%)
	4	9 (12%)
	5	8 (11%)
	6	22 (30%)
AO/OTA Classification	41-B1	18 (24%)
	41-B2	16 (21%)
	41-B3	15 (20%)
	41-C	26 (35%)
Depression (mm)		5.95 ± 4.73
Step-off (mm)		4.68 ± 3.74
Plateau widening (mm)		5.62 ± 5.84
Gap (mm)		12.93 ± 10.82

Data given as frequencies (percentages) or mean ± standard deviation.

KOOS differed significantly according to fracture severity ($p<0.001$) (Table 2). In contrast, fixation method was not associated with differences in KOOS ($p=0.817$). Additionally, fracture complexity parameters, including the number of involved columns and anterior or posterior segment counts, were not associated with significant differences in overall KOOS values ($p=0.243$, $p=0.316$, and $p=0.067$, respectively) (Table 2). Representative preoperative, postoperative, and follow-up radiographs of a patient treated with plate-and-screw fixation are presented in Figure 2.

Knee flexion also improved significantly over time; however, no significant differences were observed between groups. Flexion values did not differ significantly according to Schatzker classification ($p=0.160$), AO/OTA classification ($p=0.218$), fixation method ($p=0.882$), or fracture complexity parameters ($p>0.05$ for all comparisons) (Table 3).

In the multivariable robust regression analysis, posterior segment count (β=2.011, 95% CI: 0.306–3.716, $p=0.022$) and male sex (β=4.146, 95% CI: 0.160–8.131, $p=0.042$) were positively associated with KOOS improvement. The presence of comorbidities (β=−7.245, 95% CI: −12.852 to −1.638, $p=0.012$) and the use of external fixation before definitive surgery (β=−8.683, 95% CI: −13.325 to −4.041, $p<0.001$) were associated with less improvement in KOOS (Table 4).

Discussion

In this study, both patient-reported outcomes and objective knee motion demonstrated significant improvement over time following surgical treatment of tibial plateau fractures. However, important differences emerged when outcomes were analyzed in relation to fracture characteristics. Both the AO/OTA and Schatzker classifications were significantly associated with KOOS scores ($p<0.001$ for both), whereas neither classification showed a statistically significant association with knee flexion ($p=0.218$ and $p=0.160$, respectively). In contrast, fixation method and fracture complexity, including column and segment involvement, were not associated with either functional or objective outcomes. These findings suggest that recovery after tibial plateau fractures is multifactorial and not solely determined by fracture morphology or surgical technique.

Table 2. KOOS outcomes.

Variables		1 month	3 months	6 months	1 year	2 years	p-values
Total KOOS Scores		23.15 ± 5.87	38.13 ± 11.96	53.32 ± 13.33	67.72 ± 14.36	84.19 ± 8.74	**<0.001**
Schatzker Classification	1	25.08±6.20 ABC, a	42.38±13.31 AB, a	56.54±12.67 AB, b	72.46±14.02 AB, c	86.31±9.11 AB, d	**<0.001**
	2	22.71±4.86 ABC, a	40.29±11.21 A, b	55.86±8.11 AB, c	70.71±10.37 AB, c	85.14±8.47 AB, d	
	3	26.81±5.19 A, a	44.13±9.22 A, b	59.38±12.99 A, c	74.31±11.71 A, d	89.81±8.26 A, d	
	4	24.89±4.73 AB, a	39.78±15.60 AB, a	54.00±15.96 AB, b	68.22±19.36 AB, c	86.11±9.92 AB, c	
	5	19.00±5.24 BC, a	33.63±10.39 AB, a	52.88±15.83 AB, b	64.25±15.11 AB, b	76.38±6.70 B, c	
	6	20.27±5.17 C, a	31.55±9.11 B, a	46.09±11.08 B, b	60.23±12.40 B, c	80.59±5.88 B, d	
AO/OTA Classification	41-B1	26.61±6.10 A, a	44.89±12.59 A, b	59.28±12.57 A, c	74.83±12.97 A, c	88.50±9.31 A, d	**<0.001**
	41-B2	24.25±4.70 B, a	38.38±10.41 A, a	54.06±15.02 A, b	67.63±16.35 A, b	86.75±9.37 A, c	
	41-B3	21.73±4.83 A, a	34.53±9.61 A, b	52.07±11.02 A, c	66.60±12.02 A, d	81.80±8.08 A, e	
	41-C	20.88±5.87 A, a	35.38±12.25 A, b	49.46±13.13 A, c	63.50±14.11 A, d	81.00±6.82 A, d	
3- Column	1 column	25.91±5.12	40.65±8.30	56.65±9.89	72.70±9.65	88.87±7.40	0.243
	2 columns	22.71±5.75	38.35±13.44	52.74±14.14	68.29±15.03	83.74±9.45	
	3 columns	20.76±5.83	35.05±12.84	50.52±15.10	61.43±15.86	79.71±6.48	
Fixation Method	Single plate with additional screw(s)	24.91±4.69	39.65±12.21	54.30±11.85	68.52±14.57	86.57±8.51	0.817
	Single plate	24.26±4.85	39.58±9.28	55.42±12.13	72.26±10.78	87.16±7.68	
	Multiple plates with additional screw(s)	20.17±5.60	33.39±10.82	50.17±14.05	63.61±15.90	80.44±6.78	
	Screw fixation alone	25.75±8.86	44.75±16.70	58.38±15.30	69.38±18.04	84.50±12.26	
	Multiple Plates	19.00±4.69	33.86±11.78	46.71±16.40	61.43±12.41	77.57±7.16	
Affected Posterior Segment Count	0	24.25±5.80 a	40.25±14.34 a	51.50±17.71 b	71.00±15.77 c	86.25±12.01 c	0.316
	1	26.25±5.88 a	39.08±11.22 b	58.25±16.76 c	72.92±15.89 d	88.83±9.43 e	
	2	23.96±4.80 a	38.96±10.70 b	53.08±10.01 c	69.58±11.58 d	86.00±7.98 e	
	3	22.67±5.98 a	37.50±13.09 b	52.63±13.10 c	65.71±13.94 d	82.25±7.72 e	
	4	17.75±7.93 a	35.75±16.15 b	45.75±14.24 c	56.75±21.96 c	78.50±10.85 d	
	5	19.14±5.15 a	36.00±13.64 b	53.43±16.84 c	63.71±16.37 d	78.71±6.73 e	
Affected Anterior Segment Count	0	23.88±5.59 a	41.38±11.61 b	57.75±10.66 c	73.25±10.25 d	85.88±8.10 e	0.067
	1	25.20±5.36 a	38.80±10.47 a	53.20±11.65 b	73.20±13.41 c	90.20±10.11 c	
	2	24.67±5.76 a	40.77±11.02 b	56.30±11.25 c	70.97±12.08 d	86.80±8.58 e	
	3	22.80±5.43 a	35.50±13.89 b	50.30±16.12 c	63.50±17.67 d	81.60±7.69 e	
	4	19.80±6.98 a	39.80±14.13 b	56.20±15.97 c	69.40±15.44 d	80.60±9.86 d	
	5	17.71±4.89 a	29.00±5.45 b	42.14±10.40 c	54.43±7.98 d	76.71±4.89 e	

Values are presented as mean ± standard deviation. Analyses were performed using generalized estimating equations. The reported p-values represent the overall effect of the corresponding variable in the GEE model. Different uppercase letters indicate statistically significant differences between groups at the same follow-up point, while different lowercase letters indicate statistically significant differences between follow-up points within the same group. Post hoc comparisons were adjusted using the Bonferroni method. A p-value <0.05 was considered statistically significant.

Table 3. Knee flexion outcomes.

Variables		1 month	3 months	6 months	1 year	2 years	p-values
Total flexion range		74.12 ± 13.91	96.67 ± 17.21	111.20 ± 15.25	117.01 ± 16.25	120.91 ± 17.31	**<0.001**
Schatzker Classification	1	77.15±14.92 a	102.08±17.76 b	114.15±13.97 c	120.31±13.43 d	125.69±15.18 e	0.160
	2	76.00±10.57 a	96.71±10.63 b	110.29±9.14 bc	115.86±12.44 b	119.71±12.98 c	
	3	81.44±11.49 a	103.56±13.20 b	117.19±13.99 c	122.38±13.93 d	126.94±14.73 e	
	4	76.89±15.69 a	99.67±21.09 b	112.44±18.60 c	117.44±21.17 d	121.44±22.20 e	
	5	69.50±14.38 a	93.38±21.98 b	109.63±20.01 c	114.00±21.83 d	117.38±22.66 d	
	6	66.95±12.35 a	88.41±15.58 b	105.45±14.49 c	112.45±16.14 d	115.14±16.72 e	
AO/OTA Classification	41-B1	78.78±13.95 a	101.50±16.49 b	113.06±14.49 c	118.83±15.25 c	124.28±17.30 d	0.218
	41-B2	77.31±14.23 a	101.00±16.18 b	115.19±15.22 c	120.19±15.55 c	124.31±15.84 d	
	41-B3	73.80±13.22 a	94.93±17.26 b	110.73±15.65 c	117.07±17.38 d	120.73±17.95 e	
	41-C	69.12±13.13 a	91.65±17.65 b	107.73±15.62 c	113.77±17.01 d	116.58±17.81 e	
3- Column	1 column	79.04±11.75 a	102.04±14.74 b	114.35±13.62 c	120.65±13.51 d	125.00±14.54 e	0.410
	2 columns	75.42±13.91 a	98.90±16.75 b	113.00±15.16 c	118.84±16.29 d	123.10±17.28 e	
	3 columns	66.81±13.68 a	87.48±17.50 b	105.10±16.00 c	110.33±17.64 d	113.19±18.40 e	
Fixation Method	Single plate with additional screw(s)	75.91±11.80 a	98.17±13.11 b	111.78±10.43 c	117.00±12.25 d	121.04±13.07 e	0.882
	Single plate	80.47±13.04 a	103.95±14.78 b	116.26±14.57 c	122.42±14.24 d	126.11±14.15 e	
	Multiple plates with additional screw(s)	66.83±13.57 a	89.72±20.05 b	106.00±17.58 c	112.28±19.19 d	115.72±20.46 e	
	Screw fixation alone	74.13±19.28 a	93.13±23.34 b	108.13±21.42 c	114.88±22.36 d	119.75±25.25 e	
	Multiple Plates	69.71±9.34 a	93.86±15.69 b	112.43±15.59 c	117.00±17.44 c	121.00±19.24 d	
Affected Posterior Segment Count	0	79.75±15.37 a	104.25±15.41 b	116.75±17.61 bc	120.00±19.49 c	123.50±20.07 d	0.217
	1	78.92±13.51 a	105.58±14.49 b	117.92±15.98 c	124.67±16.63 d	129.83±17.78 e	
	2	76.54±9.70 a	99.54±11.35 b	112.67±9.28 c	119.50±10.73 d	123.71±11.65 e	
	3	73.17±15.61 a	94.50±19.18 b	109.79±16.25 c	114.54±17.16 d	117.92±17.99 e	
	4	63.75±18.19 a	87.00±29.20 b	103.00±26.62 cd	109.75±27.98 c	112.00±30.28 d	
	5	63.57±13.30 a	80.14±13.77 b	101.00±15.96 c	106.29±15.99 d	109.86±16.38 e	
Affected Anterior Segment Count	0	72.88±6.42 a	97.63±12.24 b	112.38±9.53 c	119.13±11.01 d	124.00±14.00 e	0.557
	1	77.40±9.40 a	101.60±8.38 b	111.40±8.29 c	117.80±8.76 d	124.60±11.67 e	
	2	77.53±13.81 a	99.80±16.66 b	114.03±14.87 c	119.47±15.66 d	123.20±15.97 e	
	3	73.70±15.63 a	95.70±20.02 b	109.85±17.41 c	115.60±18.66 d	119.15±19.86 e	
	4	70.20±15.85 a	93.40±21.03 b	108.60±17.78 bc	114.80±17.81 c	118.00±19.17 d	
	5	62.57±13.31 a	83.71±15.73 b	103.29±18.71 c	109.14±21.07 d	112.00±22.08 e	

Values are presented as mean ± standard deviation. Analyses were performed using generalized estimating equations. The reported p-values represent the overall effect of the corresponding variable in the GEE model. Different uppercase letters indicate statistically significant differences between groups at the same follow-up point, while different lowercase letters indicate statistically significant differences between follow-up points within the same group. Post hoc comparisons were adjusted using the Bonferroni method. A p-value <0.05 was considered statistically significant.

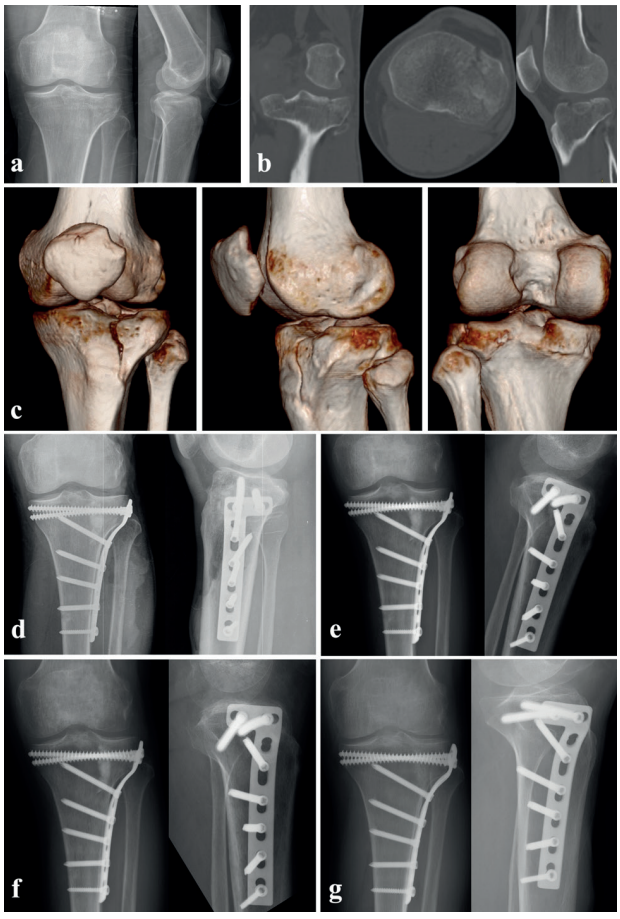


Figure 2. A 51-year-old male patient sustained a left tibial plateau fracture (Schatzker type 2, AO/OTA 41-B3) following a pedestrian traffic accident.

(a) Anteroposterior and lateral radiographs obtained at emergency department presentation. (b, c) Preoperative computed tomography (CT) images in coronal, axial, and sagittal planes and three-dimensional CT reconstructions demonstrating the fracture pattern. (d) Immediate postoperative radiographs showing reduction and fixation with a proximal tibial plate and screws, with use of allograft. (e) One-month, (f) six-month, and (g) two-year anterior-posterior and lateral follow-up radiographs demonstrating maintained reduction, fracture healing, and stable fixation over time.

The progressive improvement observed in both KOOS and knee flexion over time is consistent with previous literature. Lang et al. 2025 demonstrated that longer follow-up duration is associated with significantly better patient-reported outcomes, reflecting ongoing biological healing and functional adaptation [10]. Similarly, several longitudinal cohort studies have shown that functional recovery continues beyond the early postoperative period, often extending up to two years and even up to five years in some domains [13,14]. However, certain basic outcomes such as range of motion may plateau earlier, typically between 6 and 12 months [15]. Our findings support this pattern, emphasizing that recovery following tibial plateau fractures follows a gradual and domain-dependent trajectory rather than reaching an early universal plateau.

Fracture severity remains an important determinant of outcome; however, its effect appears to be domain specific. In the present study, KOOS scores differed significantly according to both the Schatzker and AO/OTA classifications, whereas knee flexion did not differ significantly between classification groups. This is consistent with previous reports showing that complex fracture patterns (Schatzker IV–VI) are associated with worse patient-reported outcomes, particularly in KOOS sports and QoL domains [10]. However, objective range of motion may remain comparable across fracture severities, as demonstrated in mid-term follow-up cohorts reporting no significant differences in knee flexion between simple and complex fractures [16]. In addition, Madi et al. demonstrated that greater postoperative articular step-off is associated with worse patient-reported outcomes, while Sadighi et al. reported that increased depression correlates with inferior clinical scores, with minimal impact on ROM [5,17]. Our findings extend this knowledge by demonstrating that the impact of fracture severity may be more pronounced in patient perception than in objective range of motion, suggesting that different outcome domains capture distinct aspects of recovery.

The discrepancy between patient-reported outcomes and objective knee motion represents one of the most important findings of this study. While KOOS scores were influenced by fracture severity, knee flexion was not, indicating that subjective and objective measures of recovery do not necessarily align. This observation is supported by previous literature. Elsoe et al. found that MRI-verified soft tissue injuries did not significantly affect KOOS scores at one year, highlighting the complex and multifactorial nature of patient-reported outcomes [18]. Likewise, Vaartjes et al. demonstrated in a large multicenter cohort that although many patients regain functional independence in pain and ADL domains, deficits in KOOS sports and QoL often persist, particularly in higher-demand activities [3]. These findings suggest that pain, perceived instability, activity modification, and psychosocial factors play a critical role in subjective recovery. Furthermore, mental health status has been shown to significantly influence

Table 4. Multivariable robust regression analysis of independent predictors of change in KOOS score.

Variables		β_1 (95% CI)	SE	β_2	t	p
Constant		15.812 (-1.377 - 33.001)	8.558	---	1.848	0.071
Age at surgery (year)		-0.091 (-0.214 - 0.033)	0.062	-0.185	-1.472	0.147
Male Sex		4.146 (0.16 - 8.131)	1.984	0.230	2.089	**0.042**
Presence of comorbidities		-7.245 (-12.852 - -1.638)	2.792	-0.328	-2.595	**0.012**
Articular surface depression (mm)		0.143 (-0.628 - 0.914)	0.384	0.086	0.373	0.710
Articular step-off (mm)		0.155 (-0.883 - 1.194)	0.517	0.073	0.301	0.765
Absolute plateau widening (mm)		0.167 (-0.168 - 0.502)	0.167	0.123	1.003	0.321
Fracture gap (mm)		0.113 (-0.157 - 0.382)	0.134	0.147	0.839	0.405
Fracture varus/valgus angle (°)		-0.14 (-0.31 - 0.03)	0.084	-0.206	-1.658	0.104
Posterior tibial slope (°)		0.092 (-0.037 - 0.221)	0.064	0.180	1.435	0.158
Number of involved columns (3-column concept)		0.952 (-2.197 - 4.1)	1.568	0.092	0.607	0.547
Number of involved posterior segments		2.011 (0.306 - 3.716)	0.849	0.311	2.369	**0.022**
Number of involved anterior segments		-0.148 (-2.069 - 1.774)	0.957	-0.025	-0.155	0.878
Fixation method (reference: Single Plate)	Single plate with additional screw(s)	-1.139 (-6.103 - 3.824)	2.471	-0.066	-0.461	0.647
	Multiple Plates	-4.625 (-12.333 - 3.082)	3.837	-0.173	-1.205	0.234
	Multiple plates with additional screw(s)	-2.462 (-8.39- 3.466)	2.952	-0.133	-0.834	0.408
	Screw fixation alone	3.415 (-2.879- 9.708)	3.134	0.133	1.090	0.281
Graft use (reference: none)	Allograft	0.12 (-4.75 - 4.99)	2.425	0.007	0.050	0.961
	Autograft	-5.504 (-11.04 - 0.033)	2.756	-0.257	-1.997	0.051
External fixation occurred before definitive surgery		-8.683 (-13.325- -4.041)	2.311	-0.443	-3.757	**<0.001**
Reoperation occurred		-5.724 (-13.163 - 1.715)	3.704	-0.190	-1.546	0.129

β_1 represents the unstandardized regression coefficient with 95% confidence interval (CI), and β_2 represents the standardized regression coefficient. Analyses were performed using robust multivariable linear regression. Standard errors (SE), t-values, and p-values are reported for each predictor. Categorical variables were entered into the model using dummy coding, with the specified reference categories. A p-value < 0.05 was considered statistically significant.

long-term outcomes after tibial plateau fractures [19]. Therefore, evaluation of outcomes after tibial plateau fractures should incorporate both subjective and objective measures to provide a comprehensive assessment.

Another notable finding of this study is that no statistically significant association between fixation method and clinical outcomes was identified in the present cohort. This aligns with current evidence suggesting that surgical technique may be less important than achieving adequate reduction and stability. A recent meta-analysis by Tay et al. found no significant differences in functional outcomes or complication rates between arthroscopy-assisted and conventional open reduction techniques [20]. While some studies, such as van der Linden et al., have suggested that greater fracture extent may be associated with worse patient-reported outcomes, the overall prognostic value of classification systems remains limited [21]. This is further supported by systematic reviews highlighting the variability and moderate reliability of commonly used classification systems, as well as their limited ability to predict clinical outcomes [22]. Taken together, these findings suggest that although classification systems remain valuable for describing fracture morphology and guiding surgical planning, their additional prognostic value remains debatable.

The regression analysis in this study further supports the multifactorial nature of recovery. Comorbidities were associated with worse early functional

improvement, which is consistent with existing literature demonstrating the negative impact of patient-related factors on recovery. Studies evaluating clinical outcomes and quality of life after tibial plateau fractures have shown that factors such as increased body weight and general health status negatively influence both physical and psychological recovery [23]. Additionally, injury-related factors such as polytrauma and fracture complexity have been associated with reduced range of motion at one year [4]. Interestingly, posterior segment involvement was associated with greater early improvement in our cohort; however, this finding contrasts with some previous reports suggesting worse high-demand functional outcomes and limited sports participation in such cases, often due to pain and fear of reinjury [24]. Therefore, this result should be interpreted cautiously. The use of external fixation before definitive surgery was also associated with less early improvement in KOOS. However, external fixation is generally used in patients with more severe injuries or greater soft-tissue compromise. Therefore, this association may reflect the initial injury severity rather than an independent adverse effect of external fixation itself. Overall, these findings underscore the importance of considering both patient-related and injury-related factors when evaluating prognosis and reinforce the need for a multidimensional assessment of recovery.

This study has several limitations. First, its retrospective design introduces the potential for selection bias. Second, all radiographic measurements were performed by a single observer, and interobserver reliability was not assessed. Third, stratifying our 75 patients into multiple classification and fixation categories created small subgroups. This may have introduced statistical power constraints and a risk of Type II errors. Fourth, the cohort was heterogeneous with respect to fracture pattern, associated injuries and treatment, and the fixation method was selected by the treating surgeon according to fracture morphology and soft-tissue condition rather than at random; residual confounding by indication therefore cannot be excluded. Fifth, because several subgroups were small, some statistically significant associations may not be clinically meaningful and should be regarded as exploratory; larger, preferably prospective studies are needed to confirm these findings.

Conclusion

Functional recovery after surgically treated tibial plateau fractures continued throughout the two-year follow-up period. Fracture severity was associated with patient-reported KOOS outcomes but not with knee flexion, statistically. No statistically significant associations between the evaluated fixation methods or fracture-complexity parameters and functional outcomes were identified in this cohort. These findings support the combined use of patient-reported and objective measures when evaluating recovery after tibial plateau fractures.

Author contributions

Conception: G.A., B.Y., O.M.K., U.C.K.; Design: G.A.; Data acquisition: B.Y., O.M.K., U.C.K.; Data analysis: B.Y., O.M.K.; Data interpretation: B.Y., O.M.K.; Drafting of the manuscript: B.Y., O.M.K., U.C.K.; Critical revision of the manuscript: G.A., U.C.K. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Hacettepe Üniversitesi Sağlik Bilimleri Araştirma Etik (Date: 22 July 2025, IRB No: SBA 25/682). Informed consent was obtained from all participants involved in this study.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

References

[1] Prat-Fabregat S, Camacho-Carrasco P. Treatment strategy for tibial plateau fractures: an update. EFORT Open Rev 2017;1(5):225-32. **[Crossref]**

[2] Schatzker J, Kfuri M. Revisiting the management of tibial plateau fractures. Injury 2022;53(6):2207-18. **[Crossref]**

[3] Vaartjes TP, van der Sluis FJ, Bosma E, et al. Most patients with tibial plateau fractures regain function for daily activities but do not return to their pre-injury level of sport: comparison of multicenter cohort of 1101 patients and age-related peers. Eur J Trauma Emerg Surg 2025;51(1):285. **[Crossref]**

[4] Li J, Zhang J, Zhao K, et al. Incidence and risk factors for decreased range of motion of the knee joint after surgery for closed tibial plateau fracture in adults. J Orthop Surg Res 2021;16(1):549. **[Crossref]**

[5] Madi S, Chotta S, Fadel PA, Fares MY, Zeichen J. Association between articular surface depression and functional outcomes after ORIF of tibial plateau fractures: a retrospective cohort study. Eur J Trauma Emerg Surg 2025;51(1):357. **[Crossref]**

[6] Urruela AM, Davidovitch R, Karia R, Khurana S, Egol KA. Results following operative treatment of tibial plateau fractures. J Knee Surg 2013;26(3):161-5. **[Crossref]**

[7] Gahlot N, Kunal K, Elhence A, et al. Ten-segment classification has lowest inter/intra-observer reliability as compared to schatzker, three-column and ao systems for tibial plateau fractures: a comparison based on surgeons' experience. Arch Bone Jt Surg 2023;11(4):256-61. **[Crossref]**

[8] Koksal A, Abul K, Kargın D, Cimen O, Mert M, Keskin A. An evaluation of intra and inter observer reliability of the five used classification systems of tibial plateau fractures. Acta Orthop Belg 2022;88(4):805-10. **[Crossref]**

[9] Vosoughi F, Menbari Oskouie I, Rahimdoost N, Pesantez R. Decoding tibial plateau fracture classifications: a century of individualized insights in a systematic review. EFORT Open Rev 2025;10(5):316-26. **[Crossref]**

[10] Lang C, Hörmandinger C, Klaut M, et al. Clinical outcome in tibial plateau fractures improves over time: Insights from a collaborative data network. Injury 2025;56(8):112607. **[Crossref]**

[11] Paker N, Buğdaycı D, Sabırlı F, Özel S, Ersoy S. Knee injury and osteoarthritis outcome score: reliability and validation of the Turkish version. Turkiye Klinikleri J Med Sci. 2007;27(3):350-6.

[12] Roos EM, Roos HP, Lohmander LS, Ekdahl C, Beynnon BD. Knee Injury and Osteoarthritis Outcome Score (KOOS)-development of a self-administered outcome measure. J Orthop Sports Phys Ther 1998;28(2):88-96. **[Crossref]**

[13] Gonzalez LJ, Hildebrandt K, Carlock K, Konda SR, Egol KA. Patient function continues to improve over the first five years following tibial plateau fracture managed by open reduction and internal fixation. Bone Joint J 2020;102-B(5):632-7. **[Crossref]**

[14] Millar SC, Bennett K, Fraysse F, Arnold JB, Solomon LB, Thewlis D. Longitudinal changes in lower limb joint loading up to two years following tibial plateau fracture. Gait Posture 2020;78:72-9. **[Crossref]**

[15] Christiano AV, Pean CA, Kugelman DN, Konda SR, Egol KA. Function and knee range of motion plateau six months following lateral tibial plateau fractures. J Knee Surg 2020;33(5):481-5. **[Crossref]**

[16] Neidlein C, Watrinet J, Pätzold R, et al. Patient-reported outcomes following tibial plateau fractures: mid- to short-term implications for knee function and activity level. J Clin Med 2024;13(8):2327. **[Crossref]**

[17] Sadighi M, Sabaghzadeh A, Biglari F, Ebrahimpour A, Karami A, Jafari Kafiabadi M. Clinical outcomes of Schatzker type II tibial plateau fractures using joint depression morphology: a cross-sectional study. Acta Orthop Belg 2023;89(1):128-34. **[Crossref]**

[18] Elsoe R, Motahar I, Mahdi F, Larsen P. Presence of magnetic resonance imaging verified soft tissue injuries did not significantly affect the patient-reported outcome 12 months following a lateral tibial plateau fracture: a 12-month prospective cohort study of 56 patients. Knee 2020;27(2):420-7. **[Crossref]**

[19] Kugelman D, Qatu A, Haglin J, Konda S, Egol K. Impact of psychiatric illness on outcomes after operatively managed tibial plateau fractures (OTA-41). J Orthop Trauma 2018;32(6):e221-5. **[Crossref]**

[20] Tay ST, Chen MZ, Chan YS, Kuo LT. The efficacy of arthroscopy-assisted versus stand-alone open reduction and internal fixation for treating tibial plateau fracture: a systematic review and meta-analysis. BMC Musculoskelet Disord 2024;25(1):865. **[Crossref]**

[21] van der Linden LR, Vaartjes TP, Kramer CJSA, et al. What is the impact of the fracture location on patient-reported functional outcomes in patients with lateral tibial plateau fractures? Injury 2025;56(11):112720. **[Crossref]**

[22] Millar SC, Arnold JB, Thewlis D, Fraysse F, Solomon LB. A systematic literature review of tibial plateau fractures: what classifications are used and how reliable and useful are they? Injury 2018;49(3):473-90. **[Crossref]**

[23] Höller LF, Höller S, Klinger E, et al. Clinical outcomes and quality of life of patients after surgical treatment of a tibia plateau fracture. Eur J Trauma Emerg Surg 2025;51(1):319. **[Crossref]**

[24] Quintens L, Van den Berg J, Reul M, et al. Poor sporting abilities after tibial plateau fractures involving the posterior column: how can we do better? Eur J Trauma Emerg Surg 2021;47(1):201-9. **[Crossref]**

Acta Medica 2026; 57(3): 238-245 DOI: 10.32552/actamedica.2026.1315

acta medica RESEARCH ARTICLE

Comparison of serum oxidative stress biomarkers in amateur football players and sedentary men

Müslüm Gök[1], Kazım Can Aytamur[2], Mustafa Utku Sarı[2], Serkan Aksu[3], Gönül Babayiğit İrez[2], Kemal Göral[2]

[1]Department of Medical Biochemistry, Faculty of Medicine, Muğla Sıtkı Koçman University, Muğla, Türkiye
[2]Faculty of Sport Sciences, Muğla Sıtkı Koçman University, Muğla, Türkiye
[3]Department of Physiology, Faculty of Medicine, Muğla Sıtkı Koçman University, Muğla, Türkiye

Abstract

Objectives: Physical activity's impact on oxidant production and defense systems varies based on its level, duration, and intensity. This research examines lipid oxidation, protein oxidation, and thiol-based antioxidant defense markers in the blood of amateur soccer players and sedentary men.

Methods: In this cross-sectional study, we assessed 24 healthy sedentary men and 21 amateur male football players. Serum levels of malondialdehyde (MDA), total thiol, and advanced oxidation protein products (AOPP) were measured and adjusted for protein content.

Results: Football players had non-significantly lower MDA levels than sedentary controls, indicating a descriptive trend towards reduced lipid peroxidation ($p = 0.0614$). Non-significantly lower AOPP levels were also observed in the football group ($p = 0.2023$), suggesting a descriptive trend towards reduced protein oxidation. Football players had slightly but non-significantly higher total thiol levels ($p = 0.3337$), indicating a possible trend towards increased thiol-based antioxidant capacity. In conclusion, these findings suggest that individuals who exercise regularly may have a more favourable redox profile, although the differences did not reach statistical significance.

Conclusion: The oxidative stress profile of amateur football players showed favorable but statistically non-significant trends compared with sedentary men. The combined serum levels of MDA, AOPP, and total thiol suggest that regular football training may be associated with positive changes in redox balance; however, these findings should be interpreted with caution as the differences did not reach statistical significance.

Keywords: football, oxidative stress, malondialdehyde, advanced oxidation protein products, total thiol

Introduction

Oxidative stress is a biological imbalance that arises when the body generates more reactive oxygen species than its antioxidant systems can manage. Reactive oxygen species are essential for cell communication and adaptability, but excessive amounts can damage cells [1]. This redox imbalance has been linked to the pathophysiology of numerous chronic illnesses, including diabetes, metabolic dysfunction, neurodegenerative diseases, and cardiovascular diseases [2]. Therefore, assessing biomarkers related to oxidative stress is essential in disease states and in physiological situations that may alter redox homeostasis, such as regular exercise.

Corresponding author: Muslum Gök • **Email:** muslumgok@mu.edu.tr
Received: May 11, 2026 **Accepted:** June 21, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Exercise is a recognised physiological factor that influences the oxidant-antioxidant balance in a complex manner dependent on intensity. Intense exercise may temporarily increase reactive oxygen species due to heightened oxygen intake, mitochondrial activity, and the metabolic demands of muscle contractions [2]. However, regular and optimised physical activity can improve the body's antioxidant systems and reduce baseline oxidative stress [3]. This adaptive response is often attributed to repeated moderate oxidative stressors that strengthen cellular defences against future oxidative stress, a process known as exercise-induced hormesis [3]. The direction and magnitude of this response can vary depending on the type of exercise, training duration, fitness level, and individual biological characteristics [4].

Football requires significant aerobic and anaerobic energy expenditure, frequent sprints, rapid changes of direction, and periods of intense running. Football players are an appropriate group for studying exercise-induced redox alterations due to their physiological characteristics [5]. Previous research suggests that football activity, oxidative stress, and antioxidant status may be related, particularly during matches or intense training sessions [5,6]. However, as amateur football players often have more varied training, recovery periods, diet, and lifestyle outside the sport, they may differ from professional athletes. Thus, measuring oxidative stress markers in amateur football players may provide valuable information on the biochemical effects of regular training under real sporting conditions.

Malondialdehyde (MDA) is a common marker of lipid peroxidation, indicating oxidative damage within cellular membranes [7]. Compared to lipid markers, advanced oxidation protein products (AOPP) reflect oxidative changes in proteins and may provide additional information [8]. Total thiol groups, mainly derived from molecules with sulfhydryl groups such as glutathione and protein thiols, are a reliable indicator of endogenous antioxidant potential [9]. By considering lipid, protein, and thiol-dependent antioxidant defences, these three markers together may provide a more comprehensive assessment of oxidative status.

Most previous research on sports-related oxidative stress has focused on markers of lipid peroxidation or selected antioxidant enzymes [4,10,11]. There has been less emphasis on AOPP and thiol-based antioxidant parameters in amateur football players compared to other research areas. This gap is significant because training may affect different aspects of redox balance in distinct ways. Assessing MDA, AOPP, and total thiol levels together may show whether regular football training is associated with a more favorable oxidative status.

In this context, the present study aimed to compare serum oxidative stress biomarkers in age-matched sedentary males and amateur male football players. We proposed that amateur football players show lower oxidative damage markers, as indicated by MDA and AOPP levels, and higher antioxidant capacity, as indicated by total thiol levels, compared to sedentary individuals. This study aims to advance understanding of long-term exercise-related biochemical changes in amateur athletes by analyzing these characteristics collectively.

Materials and Methods

Study design and participants

This study involved 45 males, divided into two groups. Twenty-one amateur male football players, who had competed in amateur leagues over the past year and maintained consistent training, comprised the football players group. The football players had regular training during the previous year; however, detailed quantitative training variables, such as weekly training duration, intensity, and total training load, were not available for all participants. The control group consisted of 24 healthy male university students from Muğla Sıtkı Koçman University who were sedentary and had not exercised regularly in the previous six months. The International Physical Activity Questionnaire's short form (IPAQ-SF) was used to confirm sedentary status, and only those meeting the criteria for physical inactivity were included in the control group [12].

Before sample collection, all participants were informed about the study procedures and provided written consent. The study protocol was approved by the Muğla Sıtkı Koçman University Medical and Health Sciences Research Ethics Committee (Protocol No: 250140/29). To ensure that recent physical exertion did not affect oxidative stress markers, blood samples were collected at rest, at least 24 hours after any exercise.

Blood collection and serum preparation

Eight millilitres of venous blood were collected from each participant at noon and placed in a plain, anticoagulant-free tube. Fasting before sampling was not required. As recent dietary intake may influence circulating redox markers, the non-fasting sampling condition was considered when interpreting the findings and was noted as a study limitation. Serum was separated from the blood samples by centrifugation at 1500 × g within 30 minutes of collection. Serum samples were aliquoted into Eppendorf tubes and stored at −80°C until biochemical analysis. Each biochemical measurement was performed in technical triplicate for each sample.

Determination of malondialdehyde levels

A TBARS-based commercial kit (Cayman, Cat. No: 10009055) was used to measure serum malondialdehyde (MDA) [13]. In short, serum samples and standards were treated with thiobarbituric acid reagent. The reaction mixture was then incubated for 60 minutes at 95–100°C. The reaction was stopped on ice for 5 minutes, and the tubes were centrifuged at 1600 × g for 10 minutes to separate the supernatant. After transferring the supernatant to 96-well microplates, absorbance was measured at 530 nm. MDA concentrations were calculated and expressed in µmol/L, using a linear standard curve.

Determination of total thiol levels

Total thiol levels in serum were measured using Ellman's reagent, 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB), following established spectrophotometric methods [14,15]. First, serum samples were mixed with DTNB reagent and kept in the dark at room temperature for 15 minutes. Absorbance was then measured at 412 nm. Total thiol concentrations were determined using a glutathione standard curve and are reported in µmol/L.

Determination of advanced oxidation protein products

Advanced oxidation protein products (AOPP) were quantified according to the procedure described by Witko-Sarsat et al.[8]. After dilution of serum samples with potassium phosphate buffer, potassium iodide and acetic acid were added. Absorbance of the reaction mixture was measured at 340 nm using a spectrophotometer. AOPP levels were quantified using a chloramine-T standard curve and reported as µmol/L chloramine-T equivalents.

Determination of total protein levels

A bicinchoninic acid (BCA) protein assay kit, based on the microplate method, was used to measure total protein concentration [16]. Absorbance was measured at 562 nm after incubating appropriately diluted serum samples and bovine serum albumin standards with BCA reagent. Total protein was used for normalization in biochemical analysis.

Statistical analysis

GraphPad Prism version 9 was used for statistical analyses. The Shapiro–Wilk test was used to evaluate the distribution of continuous variables. The Mann–Whitney U test was applied to non-normally distributed data, and Welch's independent samples t-test was used for normally distributed data. Data are presented as mean ± SD (standard deviation). As detailed quantitative training variables, such as weekly training duration, intensity, or training load, were not available for all participants, correlation analysis between training variables and biochemical markers was not performed. A p value less than 0.05 was considered statistically significant. Effect size calculations, including Cohen's d where appropriate, supported the assessment of group differences.

Results

A total of 45 males participated in the study, consisting of 21 amateur football players and 24 sedentary individuals. Three biochemical markers were examined to quantify resting oxidative stress: total thiol levels, AOPP, and MDA. A consistent pattern was observed between the groups when these parameters were analyzed. Oxidative damage markers were generally lower, and antioxidant thiol levels were slightly higher, in football players compared to sedentary individuals.

The IPAQ-SF was used to confirm the physical activity status of the control participants before group assignment. The IPAQ-SF evaluation showed that participants in the sedentary control group were physically inactive and had not exercised regularly in the past six months. The questionnaire findings

Table 1. Demographic and anthropometric characteristics of the study groups.

	Variable	Sedentary (n= 24)	Football Players (n= 21)	p value
BMI	Min-Max	20.4-35.1	19.9-24.6	<0.0001
	Median±SEM	26.0±0.793	21.5±0.284	
	Mean±SD	26.4±3.81	21.8±1.30	
Age	Min-Max	18.0-29.0	18.0-25.0	0.1376
	Median±SEM	18.0±0.474	22.0±0.829	
	Mean±SD	21.9±2.32	20.9±3.80	

Data are presented as mean ± SD. Between-group comparisons were performed using the independent samples Welch's t test according to data distribution. BMI: body mass index; SEM: standard error of the mean; SD: standard deviation.

validated the intended division between the physically active football group and the sedentary control group. This confirmation reinforced the validity of the group comparison by minimizing the likelihood that regularly active individuals were included in the sedentary controls.

Demographic and anthropometric characteristics of the study subjects were also evaluated (Table 1). The BMI of the sedentary group was significantly higher than that of the football players (p<0.0001). Football players had a leaner body composition, with a mean BMI of 21.8 ± 1.30 kg/m^2, compared to the sedentary group's mean BMI of 26.4 ± 3.81 kg/m^2. The age distributions of the groups were similar (p = 0.1376). The mean age was 21.9 ± 2.32 years for the sedentary group and 20.9 ± 3.80 years for the football player group. These results indicate that the groups were largely similar in age, with BMI differences reflecting the expected physical profiles of regularly trained football players.

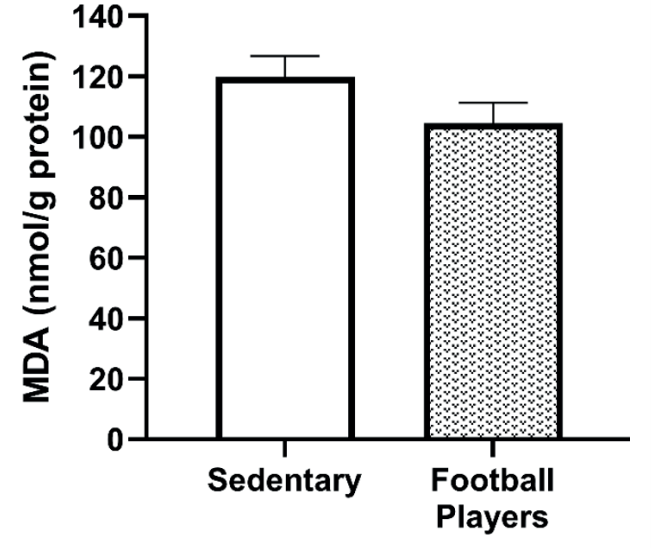


Figure 1. Serum MDA levels in amateur football players and sedentary controls. Data are presented as mean ± SEM.

MDA: malondialdehyde; SEM: standard error of the mean.

Compared with the sedentary control group, the football player group had non-significantly lower MDA levels, which indicate lipid peroxidation (Figure 1).

Sedentary individuals had a mean MDA value of 120 nmol/g protein, while football players had a mean value of 105 nmol/g protein. The football group showed a clear downward trend, although this did not reach statistical significance (p = 0.0614). The data suggest that amateur footballers may show a trend towards lower lipid peroxidation than sedentary men.

Football players also had non-significantly lower AOPP levels than sedentary individuals (Figure 2).

In the football group, the mean AOPP concentration was 2.68 µmol/g protein, compared to 3.05 µmol/g protein in the sedentary group. Although this difference was not statistically significant (p = 0.2023), the study suggests that football players tended to have lower levels of protein oxidative modification.

Total thiol levels in football players were slightly higher, in contrast to the oxidative damage markers (Figure 3).

The mean total thiol level for football players was 15.1 µmol/g protein, while the sedentary group had a mean of 14.8 µmol/g protein. The football group exhibited slightly higher thiol-dependent antioxidant capacity,

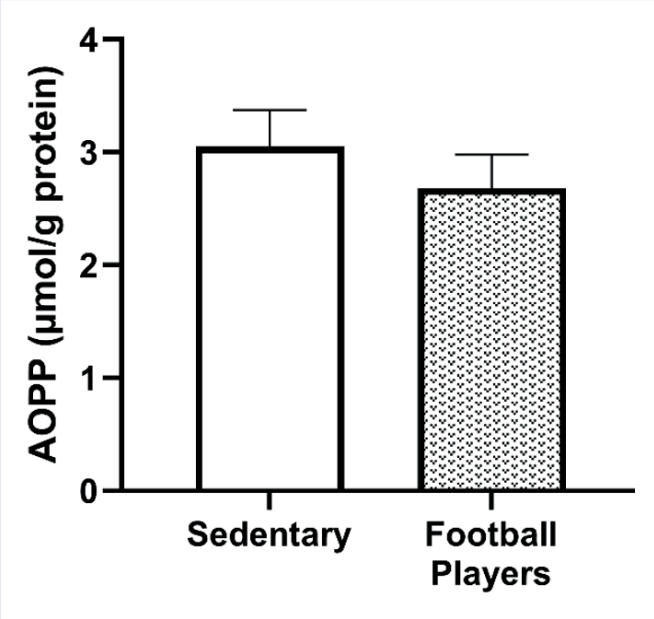


Figure 2. Serum AOPP levels in amateur football players and sedentary controls. Data are presented as mean ± SEM.

SEM: standard error of the mean.

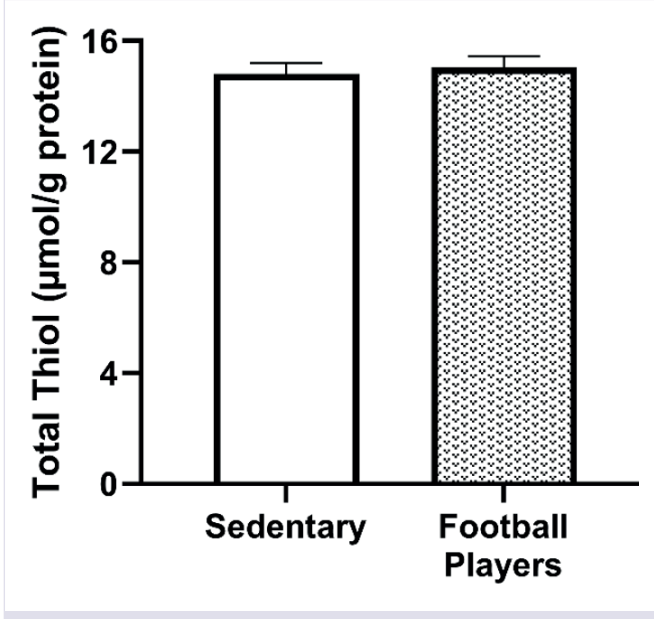


Figure 3. Serum total thiol levels in amateur football players and sedentary controls. Data are presented as mean ± SEM.

SEM: standard error of the mean.

although this was not significant (p = 0.3337). The combination of lower MDA and AOPP levels and higher total thiol levels suggests that amateur football players may have a more favorable resting redox profile than inactive individuals. However, the data should be presented as a consistent descriptive trend rather than as statistically verified group differences, as no single comparison was statistically significant.

Discussion

The present results suggest that consistent football training may be associated with favorable trends in the biological balance of oxidants and antioxidants. This conclusion is significant because football is known to induce acute oxidative stress during competition and intense training [17]. Antioxidants can lessen the first increase in TBARS/MDA that usually occurs after competitive or hard exercise [5]. Thus, long-term adaptation and acute oxidative stress are different phases of a single physiological response rather than opposing theories [6,18].

Human studies indicate that regular training may reduce resting lipid peroxidation, which is consistent with the non-significantly lower MDA levels observed in the football group. Previous research has reported that professional footballers have lower serum MDA compared with sedentary individuals [19]. Another study found that, compared to sedentary men, football players had higher total antioxidant status and lower oxidative stress index [20]. In a parallel study, football players showed lower lipid hydroperoxide and higher paraoxonase activity, demonstrating reduced oxidative stress in trained individuals [21]. In line with our MDA findings, these results suggest that regular football training may contribute to redox control; however, our findings should be interpreted with caution because the MDA difference did not reach statistical significance.

In sport studies, protein oxidation is less frequently studied than lipid oxidation, so the AOPP result provides valuable information. We found that football players exhibited non-significantly lower AOPP values, suggesting that possible redox adaptation may not be limited to lipids. This is important, as oxidative marker findings in exercise studies are frequently inconsistent. Previous research reported that intermittent jogging increased MDA but not AOPP, indicating that protein oxidation responds more slowly to short-term exercise than lipid peroxidation [18]. Another key study reported

that endurance-trained athletes showed reduced AOPP compared to strength-trained athletes, with sedentary individuals exhibiting intermediate levels [22]. This suggests that different training techniques may influence protein oxidation. Another study found that AOPP did not differ between elite sports groups, even when variations in MDA were present [11]. Together with previous studies, our AOPP results suggest a possible non-significant trend towards lower protein oxidation in football players, likely influenced by sport type and training intensity.

In this study, the slightly elevated but non-significant total thiol concentration in football players may be important, as thiol-containing compounds serve as an important extracellular antioxidant buffer. Previous research reported that frequent exercisers had greater total thiol levels than inactive participants [23]. Recent football research has shown a strong correlation between training load and thiol oxidation, indicating that thiol biology is highly responsive to exercise [24]. According to the available data, the football group exhibited a possible favorable redox trend, with slightly higher antioxidant defence markers and lower oxidative damage markers; however, this interpretation should be made cautiously, as the differences were not statistically significant.

On the other hand, there are also contrary studies. A study observed reduced erythrocyte MDA in football players compared to sedentary men [25]. However, they concluded that the athletes experienced higher oxidative stress due to lower levels of several antioxidant vitamins in that group. Similarly, Watson et al. found that inactive controls and trained athletes did not differ in a few resting oxidative stress indicators [26]. These differences are not unexpected. Results of oxidative stress are inconsistent due to variations in biological material, chosen marker, prior exercise, nutritional state, training volume for professionals and amateurs, and sample timing (during competition, recovery, or rest). Despite these results, the current dataset is internally consistent: slightly higher total thiol, lower AOPP, and lower MDA all showed the same non-significant descriptive trend. Overall, the most supportable conclusion is that consistent football training in this group may be associated with a more favorable resting redox profile compared to a sedentary lifestyle, although this should be confirmed in larger and better-controlled studies.

Future perspectives

Prospective research is needed to examine this pattern by tracking players during the preseason, the season, and their recovery. Frequent sampling may allow researchers to distinguish between temporary changes caused by match pressure or increased training adaptations. Including more enzymatic antioxidants and thiol/disulphide homeostasis indicators, such as total glutathione, in the redox panel would also be beneficial. Tracking dietary intake, body composition, aerobic fitness, sleep, and inflammatory markers may provide further valuable information. This approach would evaluate whether the current favorable descriptive trend continues throughout the season or changes with increased training intensity, as suggested by studies on training load and match recovery [5,24].

Author contributions

Conception and design: M.G., K.C.A., M.U.S.; Data acquisition: M.G., K.C.A., M.U.S., S.A., G.B.İ., K.G.; Data analysis: M.G., K.C.A.; Data interpretation: M.G., K.C.A., M.U.S., S.A., G.B.İ., K.G.; Drafting of the manuscript: M.G.; Critical revision of the manuscript: K.C.A., M.U.S., S.A., G.B.İ., K.G. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Muğla Sıtkı Koçman University Medical and Health Sciences Research Ethics Committee (Date: January 27, 2026, Decision/Protocol No: 250140/29). Informed consent was obtained from all participants involved in this study.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Acknowledgements

The authors would like to thank Ali Akbaş and Volkan Dizman for their valuable support and contributions during the study.

Generative AI statement

The authors declare that during the preparation of this study, the following AI-assisted technology was used: Not specified in the manuscript. Extent of Use: Generative AI or AI-assisted technologies were used solely for English language editing and improving the clarity of the manuscript. The authors confirm that they have critically reviewed and edited any AI-generated content and take full responsibility for the integrity, accuracy, and originality of the publication. The authors certify that the original human contribution is maintained and that AI-assisted tools are not listed or cited as authors.

References

[1] Sies H. Oxidative stress: concept and some practical aspects. Antioxidants (Basel) 2020;9(9):852. **[Crossref]**

[2] Pizzino G, Irrera N, Cucinotta M, et al. Oxidative stress: harms and benefits for human health. Oxid Med Cell Longev 2017;2017:8416763. **[Crossref]**

[3] Radak Z, Chung HY, Koltai E, Taylor AW, Goto S. Exercise, oxidative stress and hormesis. Ageing Res Rev 2008;7(1):34-42. **[Crossref]**

[4] Powers SK, Duarte J, Kavazis AN, Talbert EE. Reactive oxygen species are signalling molecules for skeletal muscle adaptation. Exp Physiol 2010;95(1):1-9. **[Crossref]**

[5] Fatouros IG, Chatzinikolaou A, Douroudos II, et al. Time-course of changes in oxidative stress and antioxidant status responses following a soccer game. J Strength Cond Res 2010;24(12):3278-86. **[Crossref]**

[6] de Oliveira DCX, Rosa FT, Simões-Ambrósio L, Jordao AA, Deminice R. Antioxidant vitamin supplementation prevents oxidative stress but does not enhance performance in young football athletes. Nutrition 2019;63-64:29-35. **[Crossref]**

[7] Ayala A, Muñoz MF, Argüelles S. Lipid peroxidation: production, metabolism, and signaling mechanisms of malondialdehyde and 4-hydroxy-2-nonenal. Oxid Med Cell Longev 2014;2014:360438. **[Crossref]**

[8] Witko-Sarsat V, Friedlander M, Capeillère-Blandin C, et al. Advanced oxidation protein products as a novel marker of oxidative stress in uremia. Kidney Int 1996;49(5):1304-13. **[Crossref]**

[9] Turell L, Radi R, Alvarez B. The thiol pool in human plasma: the central contribution of albumin to redox processes. Free Radic Biol Med 2013;65:44-253. **[Crossref]**

[10] Hammouda O, Chtourou H, Chaouachi A et al. Effect of short-term maximal exercise on biochemical markers of muscle damage, total antioxidant status, and homocysteine levels in football players. Asian J Sports Med 2012;3(4):239-46. **[Crossref]**

[11] Hadžović-Džuvo A, Valjevac A, Lepara O, Pjanić S, Hadžimuratović A, Mekić A. Oxidative stress status in elite athletes engaged in different sport disciplines. Bosn J Basic Med Sci 2014;14(2):56-62. **[Crossref]**

[12] Craig CL, Marshall AL, Sjöström M, et al. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 2003;35(8):1381-95. **[Crossref]**

[13] Ohkawa H, Ohishi N, Yagi K. Assay for lipid peroxides in animal tissues by thiobarbituric acid reaction. Anal Biochem 1979;95(2):351-8. **[Crossref]**

[14] Hu ML. Measurement of protein thiol groups and glutathione in plasma. Methods Enzymol 1994;233:380-5. **[Crossref]**

[15] Taylan E, Resmi H. The analytical performance of a microplate method for total sulfhydryl measurement in biological samples. Turk J Biochem 2010;35(3):275-8.

[16] Smith PK, Krohn RI, Hermanson GT, et al. Measurement of protein using bicinchoninic acid. Anal Biochem 1985;150(1):76-85. **[Crossref]**

[17] Tauler P, Ferrer MD, Sureda A, et al. Supplementation with an antioxidant cocktail containing coenzyme Q prevents plasma oxidative damage induced by soccer. Eur J Appl Physiol 2008;104(5):777-85. **[Crossref]**

[18] Souissi W, Bouzid MA, Farjallah MA, et al. Effect of different running exercise modalities on post-exercise oxidative stress markers in trained athletes. Int J Environ Res Public Health 2020;17(10):3729. **[Crossref]**

[19] Zanella AM, Nakazone MA, Pinhel MAS, Souza DRS. Lipid profile, apolipoprotein A-I and oxidative stress in professional footballers, sedentary individuals, and their relatives. Arq Bras Endocrinol Metabol 2011;55(2):121-6. **[Crossref]**

[20] Atli M, Aslan M, Emin Kucukoglu M, Temur HB, Taskin A, Celik H. Peripheral lymphocyte DNA damage and oxidative status in football players after a three-day football tournament. Intern Med 2013;52(2):213-7. **[Crossref]**

[21] Atli M. Serum paraoxonase activity and lipid hydroperoxide levels in adult football players after three days football tournament. Afr Health Sci 2013;13(3):565-71. **[Crossref]**

[22] Choi Y, Maeda S, Otsuki T, Miyaki A, Shimojo N, Yoshizawa M, et al. Oxidative stress and arterial stiffness in strength- and endurance-trained athletes. Artery Res 2010;4(2):52-8. **[Crossref]**

[23] Gol M, Özkaya B, Yildirim C, Bal R. Regular exercise, overweight/obesity and sedentary lifestyle cause adaptive changes in thiol-disulfide homeostasis. An Acad Bras Cienc 2019;91(2):e20180547. **[Crossref]**

[24] James C, Weber J, Boyd C, Fournier PA, Arthur PG. The thiol oxidation state of albumin is associated with training load across an Australian football pre-season. Oxid Med Cell Longev 2026;2026(1):e5534194. **[Crossref]**

[25] Siquier-Coll J, Muñoz Marín D, Grijota Pérez FJ, et al. Influence of soccer training on parameters of oxidative stress in erythrocytes. Nutr Hosp 2019;36(4):926-30.

[26] Watson TA, MacDonald-Wicks LK, Garg ML. Oxidative stress and antioxidants in athletes undertaking regular exercise training. Int J Sport Nutr Exerc Metab 2005;15(2):131-46. **[Crossref]**

Acta Medica 2026; 57(3): 246-257 DOI: 10.32552/actamedica.2026.1329

acta medica

RESEARCH ARTICLE

Distinct psychological correlates of burnout dimensions among healthcare workers

Talat Demirsöz[1], Mevhibe İrem Yıldız[1]

[1]Department of Psychiatry, Faculty of Medicine, Hacettepe University, Ankara, Türkiye

Abstract

Objective: Burnout is a multidimensional construct consisting of emotional exhaustion, depersonalization, and reduced personal accomplishment. Although psychological distress and psychological resources have both been linked to burnout, less is known about their differential associations with specific burnout dimensions. This study aimed to examine the associations of psychological symptom levels, fear of COVID-19, psychological resilience, meaning in life, and gratitude with the dimensions of burnout among healthcare workers.

Materials and Methods: A cross-sectional study was conducted with 104 healthcare workers employed at a university hospital. Participants completed self-report measures assessing burnout, psychological symptom levels, fear of COVID-19, psychological resilience, meaning in life, and gratitude. Pearson correlation analyses and hierarchical multiple regression analyses were performed.

Results: Psychological symptom severity was independently positively associated with emotional exhaustion, whereas psychological resilience was independently negatively associated with emotional exhaustion. Presence of Meaning and psychological resilience were independently negatively associated with reduced personal accomplishment. Depersonalization showed a different and more mixed pattern, and the psychological variables entered in the final step did not collectively account for a statistically significant increment in explained variance. Fear of COVID-19 was not independently associated with any of the three burnout dimensions.

Conclusion: The findings indicate that burnout dimensions may show partly distinct patterns of association with psychological and demographic characteristics. Psychological distress and lower resilience were particularly associated with emotional exhaustion, whereas lower Presence of Meaning and lower resilience were associated with reduced personal accomplishment. Findings regarding depersonalization were more mixed and should be interpreted cautiously. These results support examining burnout dimensions separately and warrant confirmation in larger and longitudinal studies.

Keywords: burnout, healthcare workers, psychological resilience, meaning in life, psychological distress

Introduction

Healthcare workers may be confronted with intense professional and psychological demands during pandemics, natural disasters, and other extraordinary circumstances that place strain on healthcare systems. Such periods of crisis may adversely affect the mental health of healthcare workers and have been associated with increased burnout [1]. During the COVID-19 pandemic, numerous studies were conducted on the mental health of healthcare workers. Specifically, depression, anxiety, sleep problems, and other indicators of psychological distress were examined extensively [2,3]. Nevertheless, understanding the

Corresponding author: Talat Demirsöz • **Email:** talatdemirsoz@gmail.com

Received: June 22, 2026 **Accepted:** September 7, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

factors associated with psychological functioning among healthcare workers continues to be an important area of inquiry beyond the pandemic.

Burnout is defined as a psychological response to chronic work-related stressors and is particularly prevalent among human service professionals [4,5]. According to the Maslach model, burnout consists of three dimensions: emotional exhaustion, depersonalization, and reduced personal accomplishment. Emotional exhaustion refers to feeling that one's emotional resources have been depleted, whereas depersonalization reflects the development of detached and impersonal attitudes toward those receiving care. The personal accomplishment dimension encompasses individuals' evaluations of their professional competence and effectiveness. Although these dimensions are conceptualized as components of the same construct in the Maslach model, research suggests that the subdimensions of burnout may be associated with different antecedents and outcomes.

Psychological symptom levels have been consistently associated with burnout among healthcare workers. The association between general psychological symptom levels and burnout has been demonstrated in various studies and systematic reviews. This association has been consistently reported, particularly with regard to depressive symptoms, anxiety symptoms, and indicators of general psychological distress [6,7]. With the emergence of the COVID-19 pandemic, researchers began to examine not only general psychological distress but also the effects of pandemic-specific threat perceptions on the mental health of healthcare workers. In this context, fear of COVID-19 has been considered one of the important variables associated with mental health and psychological functioning [8].

In addition to psychological distress, individual psychological resources may be relevant to burnout. Within the Job Demands–Resources framework, personal resources have been conceptualized as characteristics that may facilitate adaptation to occupational demands [9-11]. In the present study, three domains of psychological resources were examined: psychological resilience, meaning in life, and gratitude. These constructs represent distinct psychological characteristics that may be relevant to adaptation under stressful circumstances. Psychological resilience reflects the capacity to recover from adversity, meaning in life reflects the integration of experiences within a framework of purpose and values, whereas gratitude reflects the recognition and appreciation of positive experiences and sources of support.

Psychological resilience refers to an individual's capacity to adapt to stress and adversity and to recover from challenging experiences [12]. Meaning in life reflects the extent to which individuals perceive their lives as meaningful, purposeful, and coherent [13]. Gratitude is defined as the tendency to recognize and appreciate positive experiences and sources of support in one's life [14]. Although these variables represent different domains of psychological resources, each has been shown to be associated with psychological well-being and mental health. While the associations of psychological resilience and meaning in life with burnout among healthcare workers have been demonstrated in various studies, gratitude has been associated with psychological well-being and other positive psychological indicators [14] and has also been considered a potential protective factor against burnout among healthcare professionals [15].

The existing literature indicates that both indicators of psychological distress and psychological resources are associated with burnout. However, it remains unclear whether psychological distress and psychological resources show similar or different associations with specific dimensions of burnout. In particular, studies that simultaneously examine psychological distress indicators and psychological resources within the same model and compare their associations with different dimensions of burnout remain relatively limited. Therefore, the present study aimed to examine the associations of psychological symptom levels, fear of COVID-19, psychological resilience, meaning in life, and gratitude with emotional exhaustion, depersonalization, and reduced personal accomplishment among healthcare workers.

Based on the existing literature, we hypothesized that higher levels of psychological distress and fear of COVID-19 would be associated with higher levels of burnout, whereas greater psychological resilience, meaning in life, and gratitude would be associated with lower levels of burnout.

Materials and Methods

Study design

This study was conducted as a cross-sectional correlational study aiming to examine the associations of psychological symptom levels, fear of COVID-19, psychological resilience, meaning in life, and gratitude with different dimensions of burnout among healthcare workers. Data were collected online using self-report measures.

Participants

The study sample consisted of 104 healthcare workers employed at a university hospital. Participants were recruited through online data collection procedures, and the sample was formed using convenience sampling. Participation was voluntary, and informed consent was obtained from all participants. The study was approved by the Hacettepe University Non-Interventional Clinical Research Ethics Committee (GO 21/1019; December 21, 2021). Demographic characteristics of the participants are shown in Table 1.

Note. Values are presented as n (%) for categorical variables and mean ± standard deviation for age. Percentages were calculated based on available data; one participant had missing information on physical illness.

Measures

Maslach burnout inventory

The Maslach Burnout Inventory, developed by Maslach and Jackson, was used to assess participants' levels of burnout [4]. Normative data for Turkish healthcare personnel were reported by Ergin [16]. The scale consists of three subscales: emotional exhaustion, depersonalization, and personal accomplishment. In the present study, the three subscales were analyzed separately. Scores on the personal accomplishment subscale were coded such that higher scores indicated reduced personal accomplishment, reflecting higher levels of burnout in this dimension. Accordingly, this outcome is referred to as reduced personal accomplishment in the analyses and reporting of the present study. In the present sample, Cronbach's alpha coefficients were .88 for emotional exhaustion, .70 for depersonalization, and .77 for reduced personal accomplishment.

Brief symptom measure-25

Brief Symptom Measure-25 (BSM-25) was used to assess participants' psychological symptom levels. The scale is a self-report measure that assesses psychological symptoms experienced during the recent period. The Turkish version of the scale has been shown to provide a reliable and valid assessment of psychological symptom levels [17]. In the present study, the total score, representing participants' overall level of psychological distress, was used. The Cronbach's alpha coefficient in the present sample was .95.

Fear of COVID-19 scale

The Fear of COVID-19 Scale, developed by Ahorsu et al., was used to assess participants' levels of fear related to COVID-19 [8]. The Turkish adaptation of the scale was conducted by Satici et al.[18]. The scale consists of seven items and assesses individuals' emotional fear responses to COVID-19. The Turkish version of the scale has demonstrated adequate validity and reliability. The Cronbach's alpha coefficient in the present sample was .85.

Brief resilience scale

The Brief Resilience Scale, developed by Smith et al., was used to assess psychological resilience [12]. The Turkish adaptation of the scale was conducted by Doğan [19]. The scale consists of six items and assesses an individual's capacity to recover from stress and adversity. The Cronbach's alpha coefficient in the present sample was .90.

Meaning in life questionnaire

The Meaning in Life Questionnaire, developed by Steger et al., was used to assess meaning in life [13]. The Turkish adaptation of the scale was conducted by Demirbaş [20]. The scale consists of two subscales: Presence of Meaning and Search for Meaning. The Presence

Table 1. Sociodemographic characteristics of participants (N = 104).

Variable	n (%) / M ± SD	Min–max
Age, years	42.71 ± 9.21	25–65
Gender		
Female	69 (66.3)	
Male	35 (33.7)	
Professional group		
Physician	44 (42.3)	
Nurse	26 (25.0)	
Other healthcare workers	34 (32.7)	
Marital status		
Single	26 (25.0)	
Married	78 (75.0)	
Education level		
High school	5 (4.8)	
Associate degree	7 (6.7)	
Bachelor's degree	33 (31.7)	
Graduate degree or higher	59 (56.7)	
Physical illness		
No	79 (76.7)	
Yes	24 (23.3)	
Receiving psychological help due to mental health problems		
No	91 (87.5)	
Yes	13 (12.5)	

of Meaning and Search for Meaning subscales were analyzed separately. The Cronbach's alpha coefficients in the present sample were .84 for Presence of Meaning and .89 for Search for Meaning.

Gratitude questionnaire (GQ-6)

The Gratitude Questionnaire, developed by McCullough et al., was used to assess gratitude [14]. The Turkish adaptation of the scale was conducted by Yüksel and Oğuz Duran [21]. The scale assesses individuals' tendencies to recognize and appreciate positive experiences and sources of support in their lives. The Cronbach's alpha coefficient in the present sample was .64.

Statistical analysis

Data were analyzed using IBM SPSS Statistics version 25. Descriptive statistics were calculated for all study variables, and relationships among variables were examined using Pearson correlation analysis. Separate hierarchical multiple regression analyses were conducted for emotional exhaustion, depersonalization, and reduced personal accomplishment. Age and gender were entered in Step 1 as demographic covariates.

Age and gender were retained as demographic covariates because they are among the demographic characteristics most frequently examined in the healthcare-worker burnout literature [22]. Psychological symptom severity and fear of COVID-19 were entered in Step 2, followed by gratitude, Presence of Meaning, Search for Meaning, and psychological

resilience in Step 3. Thus, the final regression models included eight predictor variables. The hierarchical order was specified a priori to examine whether the psychological variables of primary interest explained additional variance after accounting for demographic characteristics, general psychological distress, and fear of COVID-19. All variables included in the regression analyses had complete data; therefore, the analytic sample size was N = 104 for each regression model. One participant had missing information on physical illness, which was not included in the regression analyses.

Regression assumptions and model robustness were examined comprehensively. Residual normality was evaluated using histograms and normal P–P plots; linearity and homoscedasticity were assessed using plots of standardized residuals against standardized predicted values; independence of errors was evaluated using the Durbin–Watson statistic; and multicollinearity was assessed using tolerance and variance inflation factor (VIF) values. Potentially unusual or influential observations within the analytic sample were additionally examined using studentized deleted residuals, Mahalanobis distance, leverage values, and Cook's distance.

Given the sample size and the number of predictors in the final models, a sensitivity power analysis based on the noncentral F distribution was conducted to determine the minimum overall regression effect size (Cohen's f^2) that could be detected with 80% power under the available design [23].

The analysis was based on the final model containing eight predictors, a total sample size of N = 104, α = .05, and a target power of 80%. The corresponding numerator and denominator degrees of freedom were 8 and 95, respectively. The analysis quantified the minimum overall regression effect size (Cohen's f^2) that could be detected with 80% power under the available design.

Results

Descriptive statistics and correlation analyses

Demographic characteristics of the participants are presented in Table 1. Descriptive statistics and Pearson correlation coefficients for the variables examined in the study are shown in Table 2.

Correlation analyses indicated that emotional exhaustion was positively associated with psychological symptom severity and Search for Meaning, and negatively associated with psychological resilience, gratitude, and Presence of Meaning. Depersonalization was positively associated with psychological symptom severity and Search for Meaning, and negatively associated with psychological resilience, gratitude, and Presence of Meaning. Reduced personal accomplishment was positively associated with psychological symptom severity and Search for Meaning, and negatively associated with psychological resilience, gratitude, and Presence of Meaning. Fear of COVID-19 was positively correlated with psychological symptom severity and

Table 2. Descriptive statistics and Pearson correlations among study variables (N = 104).

Variable	M	SD	Min	Max	1	2	3	4	5	6	7	8	9
1. DP	4.88	3.34	0	14	—								
2. EE	15.62	6.57	2	36	.56**	—							
3. RPA	9.85	3.98	1	23	.34**	.41**	—						
4. BSM-25	58.76	24.82	26	132	.35**	.49**	.27**	—					
5. COVID-19 Fear	12.93	4.92	7	29	.06	.07	.04	.30**	—				
6. Gratitude	29.73	6.15	15	42	−.22*	−.25*	−.34**	−.26**	−.04	—			
7. Presence	27.11	5.36	11	35	−.28**	−.35**	−.52**	−.37**	.00	.51**	—		
8. Search	17.25	7.73	5	35	.38**	.33**	.22*	.44**	.19*	−.30**	−.50**	—	
9. Resilience	20.55	5.19	6	30	−.29**	−.46**	−.45**	−.35**	−.25*	.32**	.44**	−.45**	—

DP = Depersonalization; EE = Emotional Exhaustion; RPA = Reduced Personal Accomplishment; BSM-25 = Brief Symptom Measure-25 total score. Higher RPA scores indicate greater burnout. *p < .05. **p < .01.

Table 3. Hierarchical multiple regression analyses of factors associated with burnout dimensions.

Predictor / statistic	Step 1	Step 2	Step 3
Emotional Exhaustion			
Age, β	−.306**	−.194*	−.122
Gender, β	−.088	−.096	−.105
Psychological Symptoms, β	—	.463***	.362***
Fear of COVID-19, β	—	−.059	−.116
Gratitude, β	—	—	−.046
Presence of Meaning, β	—	—	−.036
Search for Meaning, β	—	—	.024
Psychological Resilience, β	—	—	−.278**
R^2	.110	.299	.383
Adjusted R^2	.092	.271	.331
ΔR^2	.110	.189	.084
ΔF	6.242	13.353	3.218
p for ΔF	.003	<.001	.016
Depersonalization			
Age, β	−.303**	−.250*	−.219*
Gender, β	.192*	.208*	.189*
Psychological Symptoms, β	—	.293**	.174
Fear of COVID-19, β	—	.055	.018
Gratitude, β	—	—	−.034
Presence of Meaning, β	—	—	.002
Search for Meaning, β	—	—	.234*
Psychological Resilience, β	—	—	−.074
R^2	.109	.204	.267
Adjusted R^2	.092	.172	.205
ΔR^2	.109	.095	.063
ΔF	6.190	5.909	2.028
p for ΔF	.003	.004	.097
Reduced Personal Accomplishment			
Age, β	−.240*	−.188	−.080
Gender, β	−.050	−.051	−.069
Psychological Symptoms, β	—	.228*	.056
Fear of COVID-19, β	—	−.012	−.023
Gratitude, β	—	—	−.091

β = standardized regression coefficient. Step 1 included age and gender; Step 2 additionally included psychological symptom severity and fear of COVID-19; Step 3 additionally included gratitude, Presence of Meaning, Search for Meaning, and psychological resilience. ΔR^2 = change in explained variance relative to the preceding model. Dashes indicate that the variable was not entered at the respective step. Higher reduced personal accomplishment scores indicate greater burnout. *p < .05. **p < .01. ***p < .001.

Table 3. Continued.

Predictor / statistic	Step 1	Step 2	Step 3
Presence of Meaning, β	—	—	−.392**
Search for Meaning, β	—	—	−.159
Psychological Resilience, β	—	—	−.275**
R^2	.064	.112	.361
Adjusted R^2	.046	.076	.308
ΔR^2	.064	.048	.249
ΔF	3.472	2.664	9.267
p for ΔF	.035	.075	<.001

β = standardized regression coefficient. Step 1 included age and gender; Step 2 additionally included psychological symptom severity and fear of COVID-19; Step 3 additionally included gratitude, Presence of Meaning, Search for Meaning, and psychological resilience. ΔR^2 = change in explained variance relative to the preceding model. Dashes indicate that the variable was not entered at the respective step. Higher reduced personal accomplishment scores indicate greater burnout. *p < .05. **p < .01. ***p < .001.

Search for Meaning but was not significantly associated with any of the three burnout dimensions (Table 2).

The results of the hierarchical multiple regression analyses examining the independent associations of the study variables with the three burnout dimensions are shown in Table 3.

Examination of regression diagnostics did not indicate substantial violations of the assumptions of linear regression. Residual distributions and normal P–P plots were broadly consistent with normality, while plots of standardized residuals against standardized predicted values did not indicate marked departures from linearity or homoscedasticity. Durbin–Watson statistics ranged from 1.920 to 1.987, and VIF values ranged from 1.124 to 1.833, providing no indication of problematic error dependence or multicollinearity. A small number of relatively large studentized deleted residuals were observed. However, complementary influence diagnostics, including leverage and Cook's distance, did not indicate that individual observations exerted undue influence on the regression models (maximum Cook's D = .099). The sensitivity analysis indicated that, with N = 104 and eight predictors, the final regression models had 80% power to detect an overall effect of approximately $f^2 = .156$ at α = .05.

In Step 1, age and gender accounted for 11.0% of the variance in emotional exhaustion ($R^2 = .110$, p = .003). The addition of psychological symptom severity and fear of COVID-19 in Step 2 explained an additional 18.9% of the variance ($\Delta R^2 = .189$, p < .001). In Step 3, the addition of gratitude, Presence of Meaning, Search for Meaning, and psychological resilience accounted for a further significant 8.4% of the variance ($\Delta R^2 = .084$, p = .016). The final model was statistically significant, F(8, 95) = 7.36, p < .001, explaining 38.3% of the variance in emotional exhaustion ($R^2 = .383$, adjusted $R^2 = .331$). In the final model, psychological symptom severity showed a significant positive independent association with emotional exhaustion (β = .362, p < .001), whereas psychological resilience showed a significant negative independent association (β = −.278, p = .006). The remaining variables were not statistically significant in the final model.

In Step 1, age and gender accounted for 10.9% of the variance in depersonalization ($R^2 = .109$, p = .003). The addition of psychological symptom severity and fear of COVID-19 in Step 2 explained an additional 9.5% of the variance ($\Delta R^2 = .095$, p = .004). The variables entered in Step 3 accounted for an additional 6.3% of the variance; however, this increment did not reach statistical significance ($\Delta R^2 = .063$, p = .097). The final model was statistically significant, F(8, 95) = 4.32, p < .001, explaining 26.7% of the variance in depersonalization ($R^2 = .267$, adjusted $R^2 = .205$). In the final model, age was negatively associated with depersonalization (β = −.219, p = .027), whereas gender (β = .189, p = .045) and Search for Meaning (β = .234, p = .038) showed significant positive independent associations. The remaining variables were not statistically significant in the final model.

In Step 1, age and gender accounted for 6.4% of the variance in reduced personal accomplishment (R^2 = .064, p = .035). The addition of psychological symptom severity and fear of COVID-19 in Step 2 accounted for an additional 4.8% of the variance, although this increment did not reach statistical significance (ΔR^2 = .048, p = .075). In Step 3, the addition of gratitude, Presence of Meaning, Search for Meaning, and psychological resilience accounted for a substantial and statistically significant additional 24.9% of the variance (ΔR^2 = .249, p < .001). The final model was statistically significant, F(8, 95) = 6.72, p < .001, explaining 36.1% of the variance in reduced personal accomplishment (R^2 = .361, adjusted R^2 = .308). In the final model, Presence of Meaning (β = −.392, p = .001) and psychological resilience (β = −.275, p = .008) showed significant negative independent associations with reduced personal accomplishment. The remaining variables were not statistically significant in the final model.

Discussion

In the present study, the associations of psychological symptom severity, fear of COVID-19, psychological resilience, meaning in life, and gratitude with three dimensions of burnout were examined among healthcare workers. The findings revealed partly distinct patterns of association across the burnout dimensions. Emotional exhaustion was independently associated with greater psychological symptom severity and lower psychological resilience, whereas reduced personal accomplishment was independently associated with lower Presence of Meaning and lower psychological resilience. Depersonalization showed a somewhat different pattern: younger age, male gender, and higher Search for Meaning were independently associated with depersonalization in the final model, although the variables entered in Step 3 did not collectively account for a statistically significant increment in explained variance. Taken together, these findings support examining the dimensions of burnout separately rather than treating burnout as a unidimensional construct.

One of the prominent findings of the present study was the independent positive association between psychological symptom severity and emotional exhaustion. This finding suggests that emotional exhaustion may be linked not only to work-related fatigue but also to a broader pattern of psychological distress. The observed association is consistent with previous studies and systematic reviews demonstrating substantial relationships between psychological symptoms and burnout [6,7]. Given the cross-sectional design, however, the direction of this relationship cannot be determined.

Psychological resilience was also independently and negatively associated with emotional exhaustion. Psychological resilience refers to an individual's capacity to recover from stress and adversity [12]. The present finding is consistent with previous research demonstrating inverse associations between resilience and burnout among healthcare workers [24,25]. Although resilience may theoretically represent an important psychological resource in the context of occupational stress, the cross-sectional nature of the present findings does not allow it to be interpreted as a causal protective factor.

For reduced personal accomplishment, Presence of Meaning and psychological resilience showed significant independent negative associations in the final model. Of these two significant associations, Presence of Meaning had the larger absolute standardized coefficient (β = −.392). Individuals reporting a stronger established sense of meaning in life therefore tended to report lower levels of reduced personal accomplishment. This finding is consistent with previous research linking a stronger sense of purpose and meaning in life with lower levels of burnout and psychopathology among healthcare workers [26].

One possible interpretation is that maintaining a broader sense of meaning and purpose may be associated with more favorable perceptions of professional competence and accomplishment, particularly under conditions of occupational stress. However, the present study assessed meaning in life rather than work-specific meaning. Therefore, any interpretation linking Presence of Meaning specifically to meaning derived from professional roles should be regarded as tentative.

Presence of Meaning and Search for Meaning showed different patterns in relation to reduced personal accomplishment. Whereas Presence of Meaning remained independently associated with reduced personal accomplishment in the final model, Search for Meaning did not. This difference is consistent with the conceptual distinction between Presence of Meaning and Search for Meaning [13], suggesting that these two

dimensions may relate differently to burnout. Presence of Meaning reflects the experience of an established sense of meaning and purpose, whereas Search for Meaning reflects an active orientation toward finding or deepening meaning [13].

Psychological resilience was also independently associated with lower reduced personal accomplishment. Individuals with greater resilience may be better able to maintain functioning and a sense of effectiveness when facing occupational setbacks and stressful experiences. However, longitudinal studies are needed to determine whether resilience helps maintain professional accomplishment over time.

Depersonalization showed a different pattern from the other two burnout dimensions. In the final model, younger age and male gender were independently associated with higher depersonalization, and Search for Meaning also showed a significant positive individual association. Importantly, however, the addition of the variables entered in Step 3 did not collectively produce a statistically significant increment in explained variance ($\Delta R^2 = .063$, $p = .097$). The association between Search for Meaning and depersonalization should therefore be interpreted cautiously and warrants replication before firm conclusions are drawn regarding its independent role.

The negative association between age and depersonalization may reflect processes related to professional experience or adaptation to occupational demands. Younger healthcare workers may have had less opportunity to develop strategies for managing sustained occupational stress. However, because professional experience was not directly assessed, age should not be interpreted as a proxy for experience, and the mechanisms underlying this association remain to be examined.

Male participants also reported higher depersonalization scores. This finding should be interpreted cautiously given the smaller male subgroup (n = 35) and because potential explanatory factors such as working conditions, organizational characteristics, and interpersonal coping styles were not assessed. Larger and more balanced samples are needed before firm conclusions regarding gender differences in depersonalization can be drawn.

More broadly, depersonalization is characterized by detached and emotionally withdrawn attitudes toward individuals receiving care or services [5]. Organizational and occupational factors not assessed in the present study may therefore also be relevant to this dimension. Previous research among healthcare professionals during the COVID-19 pandemic has linked perceptions of the work and educational environment with burnout, highlighting the potential importance of contextual factors [27].

Taken together, the findings indicate that the three burnout dimensions were characterized by partly distinct patterns of association. Emotional exhaustion was independently associated with psychological distress and resilience, whereas reduced personal accomplishment was associated with Presence of Meaning and resilience. Depersonalization showed a more mixed pattern involving demographic characteristics and Search for Meaning, although the incremental contribution of Step 3 was not statistically significant. These differences support examining burnout dimensions separately, as reliance solely on an overall burnout score may obscure potentially meaningful distinctions among emotional exhaustion, depersonalization, and reduced personal accomplishment.

Fear of COVID-19 was not independently associated with any of the three burnout dimensions in the final regression models. Although fear of COVID-19 was correlated with psychological symptom severity, its coefficients were not statistically significant after the other variables were considered. This pattern suggests that the unique association between COVID-19 fear and burnout was limited in the present sample. However, given the cross-sectional design and the specific context and timing of data collection, this finding should not be interpreted as evidence that pandemic-related fear is generally unrelated to burnout. The absence of an independent association between fear of COVID-19 and burnout may also have been influenced by contextual factors that were not assessed in the present study, including differences in occupational exposure and institutional working conditions. Therefore, fear of COVID-19 as an individual emotional response should be distinguished from the broader occupational and psychosocial burden associated with working during the pandemic. The heterogeneity of professional groups in our sample should also be considered when interpreting the absence of an independent association between fear of COVID-19 and burnout. The sample included physicians, nurses, and other healthcare workers, whose

levels of patient contact, occupational responsibilities, perceived infection risk, and psychosocial burden may have differed. Therefore, the absence of a significant association in the overall sample does not exclude the possibility that the relationship between COVID-19-related fear and burnout may differ across professional groups. Larger studies with sufficient power for subgroup analyses are needed to examine these potential differences.

Gratitude was significantly correlated with all three burnout dimensions but was not independently associated with them in the final regression models. This pattern may partly reflect shared variance with other psychological resources such as resilience and meaning in life. However, the relatively modest internal consistency of the Gratitude Questionnaire in the present sample (α = .64) may also have attenuated associations involving gratitude. Accordingly, the absence of significant independent associations should be interpreted cautiously rather than as evidence that gratitude is unrelated to burnout. Previous experimental and correlational research has also linked gratitude with higher psychological well-being and positive affect [14,28].

Several limitations should be considered when interpreting the present findings. First, the cross-sectional design precludes conclusions regarding the temporal or causal direction of the observed associations. Although hierarchical regression allowed us to examine associations with each burnout dimension while accounting for the other variables in the models, the findings should not be interpreted as evidence that these factors prospectively predict or causally influence burnout. Longitudinal studies are needed to clarify the temporal relationships among psychological distress, psychological resources, and burnout. Second, the relatively modest sample size should be considered when interpreting the regression findings. A sensitivity analysis indicated that the available sample (N = 104) provided 80% power to detect an overall regression effect of approximately $f^2 = .156$ in the final eight-predictor models. Thus, the study had reasonable sensitivity to approximately moderate overall effects, but smaller overall effects and, importantly, smaller unique or incremental effects may have been insufficiently powered. This consideration is particularly relevant to the depersonalization model, in which the variables entered in Step 3 accounted for an additional 6.3% of the variance but the increment did not reach statistical significance. Replication in larger samples is therefore warranted. Third, participants were recruited using convenience sampling from a single university hospital, which may limit the generalizability of the findings to healthcare professionals working in other institutions, occupational settings, or sociocultural contexts. The smaller male subgroup (n = 35) also warrants caution when interpreting the observed association between gender and depersonalization. Fourth, all study variables were assessed using self-report measures, which may have introduced common method variance and response-related biases. Moreover, the internal consistency of the Gratitude Questionnaire was relatively modest in the present sample (α = .64), which may have attenuated associations involving gratitude. Findings concerning gratitude should therefore be interpreted with particular caution. Fifth, online recruitment may have introduced selection bias, as healthcare workers who chose to participate may have differed systematically from those who did not participate. Finally, detailed occupational and organizational characteristics, such as workload, working hours, direct exposure to patients with COVID-19, and perceived organizational support, were not assessed. Such factors may account for additional variance in burnout dimensions and may be particularly relevant to depersonalization.

The present study demonstrated partly distinct patterns of association between psychological characteristics and the three dimensions of burnout among healthcare workers. Emotional exhaustion was independently associated with greater psychological symptom severity and lower psychological resilience, whereas reduced personal accomplishment was independently associated with lower Presence of Meaning and lower psychological resilience. Depersonalization showed a different and more mixed pattern involving age, gender, and Search for Meaning, although the variables entered in Step 3 did not collectively account for a statistically significant increment in explained variance. These findings highlight the value of examining burnout as a multidimensional construct rather than relying solely on an overall burnout score. Longitudinal studies with larger and more diverse samples are needed to determine the temporal nature of these associations and to clarify their potential relevance for interventions aimed at healthcare-worker well-being.

Author contributions

Conception and design: T.D., M.İ.Y.; Data acquisition: T.D., M.İ.Y.; Data analysis: T.D.; Data interpretation: T.D., M.İ.Y.; Drafting of the manuscript: T.D.; Critical revision of the manuscript: T.D., M.İ.Y. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Hacettepe University Non-Interventional Clinical Research Ethics Committee (Date: December 21, 2021, Decision/Protocol No: GO 21/1019). Informed consent was obtained from all participants involved in this study.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that during the preparation of this study, the following AI-assisted technology was used: ChatGPT on April-August 2026. Extent of Use: ChatGPT was used during manuscript preparation to assist with English language editing, improving the clarity and organization of the text, refining the presentation of statistical results, and supporting literature exploration. All cited sources were independently checked against the original publications, and all statistical analyses and numerical results were independently conducted and verified by the authors. The authors confirm that they have critically reviewed and edited any AI-generated content and take full responsibility for the integrity, accuracy, and originality of the publication. The authors certify that the original human contribution is maintained and that AI-assisted tools are not listed or cited as authors.

References

[1] Ghahramani S, Lankarani KB, Yousefi M, Heydari K, Shahabi S, Azmand S. A systematic review and meta-analysis of burnout among healthcare workers during COVID-19. Front Psychiatry 2021;12:758849. **[Crossref]**

[2] Pappa S, Ntella V, Giannakas T, Giannakoulis VG, Papoutsi E, Katsaounou P. Prevalence of depression, anxiety, and insomnia among healthcare workers during the COVID-19 pandemic: a systematic review and meta-analysis. Brain Behav Immun 2020;88:901-7. **[Crossref]**

[3] Serrano-Ripoll MJ, Meneses-Echavez JF, Ricci-Cabello I, et al. Impact of viral epidemic outbreaks on mental health of healthcare workers: a rapid systematic review and meta-analysis. J Affect Disord 2020;277:347-57. **[Crossref]**

[4] Maslach C, Jackson SE. The measurement of experienced burnout. J Occup Behav 1981;2:99-113. **[Crossref]**

[5] Maslach C, Schaufeli WB, Leiter MP. Job burnout. Annu Rev Psychol 2001;52:397-422. **[Crossref]**

[6] Koutsimani P, Montgomery A, Georganta K. The relationship between burnout, depression, and anxiety: a systematic review and meta-analysis. Front Psychol 2019;10:284. **[Crossref]**

[7] Zhu H, Yang X, Xie S, Zhou J. Prevalence of burnout and mental health problems among medical staff during the COVID-19 pandemic: a systematic review and meta-analysis. BMJ Open 2023;13(7):e061945. **[Crossref]**

[8] Ahorsu DK, Lin CY, Imani V, Saffari M, Griffiths MD, Pakpour AH. The fear of COVID-19 scale: development and initial validation. Int J Ment Health Addict 2022;20(3):1537-45. **[Crossref]**

[9] Demerouti E, Bakker AB, Nachreiner F, Schaufeli WB. The job demands-resources model of burnout. J Appl Psychol 2001;86(3):499-512.

[10] Bakker AB, Demerouti E. The job demands-resources model: state of the art. J Manag Psychol 2007;22:309-328. **[Crossref]**

[11] Xanthopoulou D, Bakker AB, Demerouti E, et al. The role of personal resources in the job demands-resources model. Int J Stress Manag 2007;14:121-141. **[Crossref]**

[12] Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008;15(3):194-200. **[Crossref]**

[13] Steger MF, Frazier P, Oishi S, et al. The meaning in life questionnaire: assessing the presence of and search for meaning in life. J Couns Psychol 2006;53:80-93. **[Crossref]**

[14] Mccullough ME, Emmons RA, Tsang JA. The grateful disposition: a conceptual and empirical topography. J Pers Soc Psychol 2002;82(1):112-27. **[Crossref]**

[15] Burke J, O'Donovan R. Gratitude as a protective factor against burnout in healthcare professionals: a systematic review. Br J Healthc Manag 2023;29(4):1-11. **[Crossref]**

[16] Ergin C. Maslach tükenmişlik ölçeğinin Türkiye sağlık personeli normları. Psikiyatri Psikoloji Psikofarmakoloji Dergisi 1996;4:28-33.

[17] Gülüm IV, Soygüt G. Psychometric properties of the Turkish brief symptom measure-25. Curr Psychol 2019;38:1558-63. **[Crossref]**

[18] Satici B, Gocet-Tekin E, Deniz ME, Satici SA. Adaptation of the fear of COVID-19 scale: its association with psychological distress and life satisfaction in Turkey. Int J Ment Health Addict 2021;19:1980-8. **[Crossref]**

[19] Doğan T. Kısa Psikolojik Sağlamlık Ölçeği'nin Türkçe uyarlaması: geçerlik ve güvenirlik çalışması. J Happiness Well-Being 2015;3:93-102.

[20] Demirbaş N. Yaşamda anlam ve yılmazlık [yüksek lisans tezi]. Ankara, Türkiye: Hacettepe Üniversitesi; 2010.

[21] Yüksel A, Oğuz Duran N. Turkish adaptation of the gratitude questionnaire. Eurasian J Educ Res 2012;46:199-216.

[22] Meredith LS, Bouskill K, Chang J, Larkin J, Motala A, Hempel S. Predictors of burnout among US healthcare providers: a systematic review. BMJ Open 2022;12(8):e054243. **[Crossref]**

[23] Faul F, Erdfelder E, Buchner A, Lang AG. Statistical power analyses using G*Power 3.1: tests for correlation and regression analyses. Behav Res Methods 2009;41(4):1149-60. **[Crossref]**

[24] Deldar K, Froutan R, Dalvand S, Gheshlagh RG, Mazloum SR. The relationship between resiliency and burnout in Iranian nurses: a systematic review and meta-analysis. Open Access Maced J Med Sci 2018;6(11):2250-6. **[Crossref]**

[25] Bellicoso D, Valenzano TJ, Santiago C, Romano D, Canzian S, Topolovec-Vranic J. Resilience and burnout among healthcare staff during COVID-19: lessons for pandemic preparedness. Healthcare (Basel) 2026;14(2):195. **[Crossref]**

[26] O'Higgins M, Rojas LA, Echeverria I, Roselló-Jiménez L, Benito A, Haro G. Burnout, psychopathology and purpose in life in healthcare workers during COVID-19 pandemic. Front Public Health 2022;10:926328. **[Crossref]**

[27] Atılgan B, Yıldız Mİ, Yavuz CI. Perceptions of work and educational environment as predictors of burnout among residents during COVID-19 pandemic. Acta Medica 2023;54(1):35-46. **[Crossref]**

[28] Emmons RA, McCullough ME. Counting blessings versus burdens: an experimental investigation of gratitude and subjective well-being in daily life. J Pers Soc Psychol 2003;84(2):377-89. **[Crossref]**

Acta Medica 2026; 57(3): 258-265 DOI: 10.32552/actamedica.2026.1350

acta medica RESEARCH ARTICLE

Comparative genomic profiling of early-onset and late-onset colorectal cancer

Bengi Öztürk[1], Hatime Arzu Yaşar[2], Onur Keskin[1], Taylan Kav[1]

[1]Department of Gastroenterology, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
[2]Department of Medical Oncology, Faculty of Medicine, Ankara University, Ankara, Türkiye

Abstract

Objective: The increasing incidence of early-onset colorectal cancer (EO-CRC) has raised the question of whether these tumors exhibit different molecular characteristics from late-onset colorectal cancer (LO-CRC). We aimed to compare the somatic genomic profiles of EO-CRC and LO-CRC.

Materials and Methods: Clinical and genomic data of patients diagnosed with colon adenocarcinoma, colorectal adenocarcinoma, and rectal adenocarcinoma from the AACR Project GENIE Cohort v20.0 database were retrospectively evaluated. Patients were divided into two groups based on age at diagnosis: early-onset (<50 years) and late-onset (≥50 years). The groups were compared in terms of demographic characteristics, number of mutations, altered genome fraction, and gene mutation frequencies. Multiple-testing correction was performed using the Benjamini–Hochberg method in the main analysis, genes with a mutation frequency ≥5% in at least one group were focused on.

Results: A total of 23,410 samples from 22,142 patients were evaluated. TP53 (74.62% vs. 69.53%), APC (70.20% vs. 65.82%), and TCF7L2 (14.61% vs. 12.27%) mutations were more frequent in the EO-CRC group, while BRAF (7.16% vs. 13.55%) and RNF43 (6.98% vs. 10.80%) mutations were more frequent in the LO-CRC group (all q<0.01). Additionally, AXIN2, PLCG2, AMER1, MSH3, MGA, SOX9, and SPEN mutations were found to be significantly more frequent in the LO-CRC group.

Conclusion: There are significant differences in somatic genomic profiles between early-onset and late-onset colorectal cancers. The more frequent occurrence of TP53, APC, and TCF7L2 mutations in EO-CRC, and BRAF and RNF43 mutations in LO-CRC, suggests that the relative contribution of molecular pathways involved in colorectal carcinogenesis may differ according to age. However, these findings do not support the evaluation of EO-CRC as a completely separate molecular disease based solely on age.

Keywords: colorectal cancer, genomic variation, early onset

Introduction

Colorectal cancer (CRC) is the third most common cancer worldwide. In terms of mortality rates, it is the second leading cause of cancer-related deaths [1]. Related to the widespread use of screening programs, the incidence and mortality of colorectal cancer detected in individuals aged 50 and over have shown a declining trend over time [2]. However, the incidence of early-onset colorectal cancer (EO-CRC), defined as colorectal cancer diagnosed before the age of 50 years, has been increasing in recent decades [1-5].

The majority of EO-CRC cases originate from the rectum and distal colon. Cases typically present with an aggressive pathology, often as poorly differentiated

Corresponding author: Bengi Öztürk • **Email:** drbengiozturk@gmail.com
Received: August 17, 2026 **Accepted:** September 10, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

carcinoma or signet ring cell carcinoma. Patients with EO-CRC are usually diagnosed at an advanced stage. All these characteristics present challenges in clinical management [3,5-8].

Hereditary predisposition plays a significant role in the development of EO-CRC. Germline pathogenic variations in cancer susceptibility genes, such as Lynch syndrome (MLH1, MSH2, MSH6, PMS2 mutations), Familial Adenomatous Polyposis (FAP) [adenomatous polyposis coli (APC) mutations], or MUTYH mutations, are detected in approximately 20% to 25% of young colorectal cancer patients compared to older patients [3,4,9,10]. However, the vast majority (75-80%) of EO-CRC cases consist of sporadic tumors that do not carry a hereditary predisposition [3,6].

This epidemiological and clinical divergence has raised the question of whether EO-CRC is a distinct entity at the biological and molecular level from late-onset colorectal cancer (LO-CRC). Furthermore, understanding the molecular mechanisms behind sporadic cases is critical for optimizing both early diagnostic detection and personalized treatment approaches.

The aim of this study is to investigate the differences in genetic variation between early-onset and late-onset colorectal cancers.

Materials and Methods

AACR project GENIE

The American Association for Cancer Research (AACR) Project Genomics, Evidence, Neoplasia, Information, and Exchange (GENIE) is an international collaborative initiative established to facilitate the sharing of genomic and clinical data and support precision oncology research [11-13]. Data analyzed in the present study were obtained from the publicly available AACR Project GENIE cohort v20.0 through cBioPortal for Cancer Genomics.

AACR Project GENIE integrates routinely generated clinical and genomic data from participating institutions worldwide. Genomic testing data contributed to the registry are generated in Clinical Laboratory Improvement Amendments (CLIA)-certified and/or International Organization for Standardization (ISO)-accredited laboratory environments [13]. Data sharing within the consortium is governed by institutional agreements and the Institutional Review Board (IRB) requirements of participating centers. Depending on institutional policies, authorization for data contribution may be based on prospective patient consent, an IRB-approved waiver, or approval of GENIE-specific research protocols [13]. Data made available for research are de-identified to protect patient confidentiality.

Statistical analyses

This study included patients diagnosed with colon adenocarcinoma, colorectal adenocarcinoma, and rectal adenocarcinoma from the AACR Project GENIE Cohort v20.0 database [13-15]. Demographic characteristics, clinicopathological findings, and genomic data of the patients were analyzed retrospectively. Patients were categorized into two groups according to age at diagnosis: early-onset colorectal cancer (EO-CRC), defined as <50 years, and late-onset colorectal cancer (LO-CRC), defined as ≥50 years. The groups were compared in terms of demographic characteristics and genetic mutations. The Chi-square test was used for categorical variables, the Kruskal-Wallis test for continuous variables, and the Wilcoxon test for comparing paired groups with abnormal distributions. The Benjamini-Hochberg method was applied to correct P-values, and consequently, false discovery rate-adjusted q-values were computed. Results were considered statistically significant when both the q-value and P-value were below 0.05. For clarity and to focus on more prevalent molecular events, genes with a mutation frequency of ≥5% in at least one group were highlighted in the main analysis.

Results

A total of 23,410 samples from 22,142 patients were evaluated. Data centers are shown in Table 1. A total of 16,667 samples from 15,858 patients belonged to patients aged 50 and over (LO-CRC). A total of 6,701 samples from 6,289 patients belonged to patients under 50 years of age (EO-CRC).

In LO-CRC group, 9,630 (57.8%) samples were taken from the primary site, while 4,877 (29.3%) samples were taken from the metastatic site. Ninety-six (0.6%) patients had cell-free DNA samples. Information on 2,064 (12.4%) samples could not be obtained. Of the patients with LO-CRC, 9,585 (60.4%) were Caucasian.

Table 1. Data center.

Data center	N (%)
Memorial Sloan Kettering Cancer Center	8228 (37.2%)
Dana-Farber Cancer Institute	4317 (19.5%)
Providence Health and Services Cancer Institute	2419 (10.9%)
The Herbert Irving Comprehensive Cancer Center, Columbia University	1349 (6.1%)
Johns Hopkins Sidney Kimmel Comprehensive Cancer Center	1347 (6.1%)
The University of Texas MD Anderson Cancer Center	640 (2.9%)
Vall d'Hebron Institute of Oncology	611 (2.8%)
Duke Cancer Institute, Duke University Health System	608(2.7%)
Netherlands Cancer Center, The Netherlands	485 (2.2%)
Princess Margaret Cancer Centre, University Health Network	466 (2.1%)
Vanderbilt-Ingram Cancer Center	438 (2.0%)
University of California-San Francisco (UCSF Helen Diller Family Comprehensive Cancer Center)	367 (1.7%)
Wake Forest University Health Sciences, Wake Forest Baptist Medical Center	278 (1.3%)
Yale University, Yale Cancer Center	181 (0.8%)
The University of Chicago	177 (0.8%)
Institut Gustave Roussy	98 (0.4%)
Swedish Cancer Institute.	97 (0.4%)
University of Miami - Sylvester Comprehensive Cancer Center	36 (0.2)

*Data are represented as numbers (N) and percentages (%).

In EO-CRC patients, 4244 (63.3%) samples were from the primary site, 1948 (29.1%) from the metastatic site. Twenty-two (0.3%) samples were from cell-free DNA. A total of 487 (7.3%) samples were of unknown origin. Of the patients with EO-CRC, 3898 (62%) were Caucasian.

A statistically significant difference was found in gender distribution between early-onset and late-onset colorectal adenocarcinoma groups ($p < 0.001$; $q < 0.001$). Male patients constituted the highest proportion in both groups, although the proportion of men was relatively higher in the late-onset group (Figure 1).

There was a statistically significant difference in mutational count distribution between the early-onset and late-onset colorectal cancer groups. The median mutational count was 8 (q1-q3:4-12) in late-onset patients, and the mutational count was 7 (q1-q3:4-11) in early-onset patients (q-value <0.001, p-value <0.001).

There was a statistically significant difference in the distribution of fractionally altered genome between the early-onset and late-onset colorectal cancer groups. The median fractional genome altered distribution was 0.17 (q1-q3: 0.05-0.31) in late-onset patients and 0.16 (q1-q3: 0.05-0.31) in early-onset patients (q-value <0.005, p-value <0.005).

Mutation frequency analysis revealed significant differences between early-onset (EO) and late-onset (LO) colorectal adenocarcinomas (Figure 2). Among statistically significant genes ($q < 0.05$), only those with a mutation frequency of ≥5% in at least one group were included in the main analysis. BRAF (13.55% vs. 7.16%; $q < 0.01$) and RNF43 (10.80% vs. 6.98%; $q < 0.01$) mutations were significantly more frequent in the LO group. In contrast, TP53 (74.62% vs. 69.53%; $q < 0.01$) and APC (70.20% vs. 65.82%; $q < 0.01$) mutations were significantly more frequent in the EO group. TCF7L2 mutations were also enriched in the EO group (14.61% vs. 12.27%; $q < 0.01$), whereas AXIN2, PLCG2, AMER1, MSH3, MGA, SOX9, and SPEN mutations were significantly enriched in the LO group (all $q < 0.05$) (Table 2).

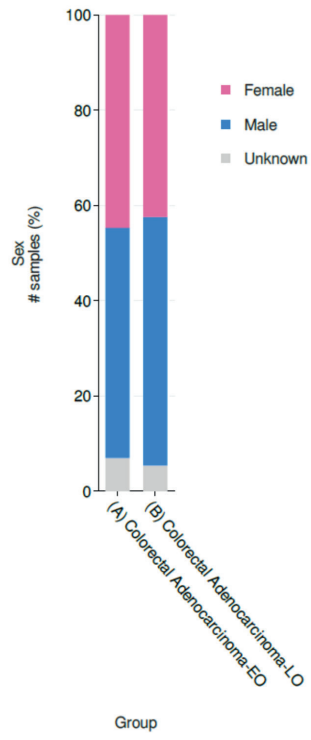


Figure 1. Sex distribution of patients with early-onset (EO) and late-onset (LO) colorectal adenocarcinoma.

Blue bars represent males (Colorectal adenocarcinoma EO, N = 3217, 48.34%; Colorectal carcinoma LO, N =8668, 52.16%), and pink bars represent females (Colorectal adenocarcinoma EO, N = 2974, 44.69%; Colorectal carcinoma LO, N =7049, 42.42%).

A significant difference in sex distribution was observed between the two groups p<0.01,q<0.01

Data are represented as numbers (N) and percentages (%).

Table 2. Differentially mutated genes in early-onset and late-onset colorectal cancer.

Gene	EO-CRC	LO-CRC	p-value	q-value
AXIN2	226 (4.65%)	759 (6.86%)	<0.01	<0.01
PLCG2	149 (4.06%)	522 (6.05%)	<0.01	<0.01
AMER1	265 (6.51%)	827 (8.62%)	<0.01	<0.01
MSH3	217 (6.20%)	704 (8.36%)	<0.01	<0.01
MGA	269 (5.83%)	775 (7.53%)	<0.01	<0.05
SOX9	593 (10.82%)	1,597 (12.78%)	<0.01	<0.05
SPEN	262 (6.48%)	792 (8.29%)	<0.01	<0.05

EO-CRC: Early onset colorectal carcinoma, LO-CRC: Late-onset colorectal carcinoma.

Discussion

In this study, we retrospectively demonstrated differences in genomic mutations between patients with early-onset and late-onset colorectal cancer in a large cohort. In our study, APC and TP53 mutations were more frequently observed in early-onset colorectal carcinoma, while BRAF and RNF43 mutations were more common in late-onset colorectal carcinoma. Less frequently, mutations such as AXIN2, PLCG2, AMER1, MSH3, MGA, SOX9, and SPEN were more common in late-onset colorectal carcinoma.

The increasing incidence of early-onset colorectal carcinoma in recent years and the aggressive course of this disease have led to increased research on this patient group.

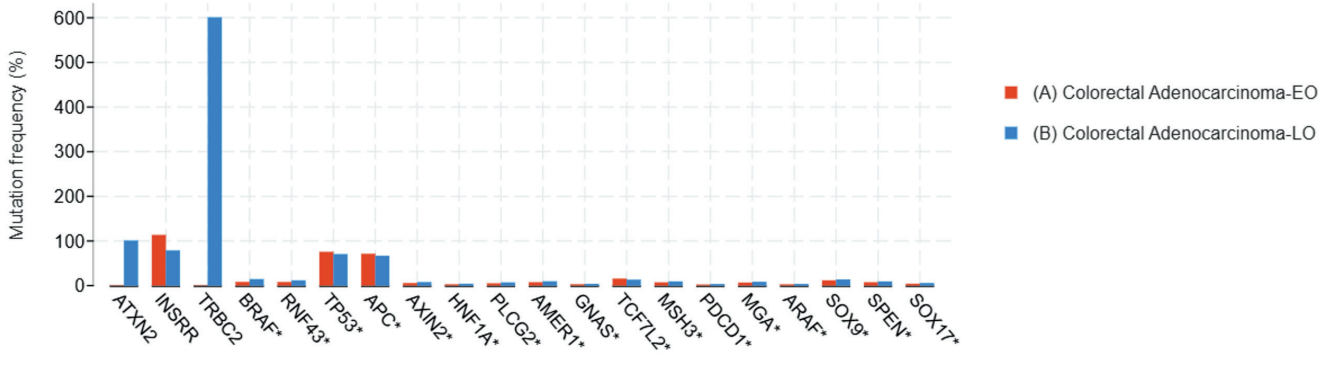


Figure 2. Gene mutations with the most significant value.

A) Blue bars represent Colorectal adenocarcinoma Late onset (LO), and (B) red bars represent. Colorectal adenocarcinoma Early onset (EO)

Data are represented as percentages (%).

*The difference between the groups was statistically significant (q < 0.05).

TP53 mutation have an important regulatory effect on colorectal tumor formation, particularly during the adenoma-carcinoma transition. TP53 mutation lead to various functional changes affecting invasion, metastasis, apoptosis, and drug resistance [16]. TP53 mutation is one of the most common mutations in patients with colorectal cancer [17,18]. Lieu et al. showed that TP53 alterations are more frequently observed in patients diagnosed with EO-CRC [19]. Li et al. showed that the TP53 mutation was found more frequently in patients with non-hypermutated colon cancer than in patients diagnosed with EO-CRC [6]. Lawler et al. reported in a meta-analysis of 149 articles that TP53 mutation is more frequently observed in EO-CRC [20]. In the study by Pretta et al., which evaluated 1209 colorectal cancer patients using Foundation One, it was also reported that TP53 mutation was more common in EO-CRC patients [21]. In our study, similar to the literature, TP53 mutation was found to be more prevalent in EO-CRC patients.

Inactivation of APC, a tumor suppressor gene, has been considered a significant early event in up to 85% of sporadic colorectal cancers [22,23]. APC is one of the key tumor suppressor genes in colorectal carcinogenesis and plays a critical role in the regulation of the WNT signaling pathway. The central role of APC in both sporadic CRC and inherited syndromes such as FAP is well understood [10,22-25]. However, previous studies examining the distribution of APC mutations between young and elderly patients have reported heterogeneous results. Some studies have reported lower rates of APC mutations in EO-CRC, particularly in specific molecular subgroups, while others have shown similar or higher frequencies in EO-CRC [6,7,17-19,21]. In our study we showed that APC mutations were significantly more frequent in the EO-CRC group. TCF7L2 is one of the most frequently mutated genes in colorectal cancer. It is a key transcriptional effector of the WNT signaling pathway [26]. In our study, we showed that TCF7L2 mutations were more frequent in the EO-CRC group. In the study by Li et al., TCF7L2 was also observed more frequently in early-onset colorectal cancer with a high tumor mutation burden (TMB > 15 mutations/Mb) [6]. The higher frequency of APC and TCF7L2 alterations in the same age group suggests potential age-related differences in the contribution of WNT pathway alterations to colorectal carcinogenesis. Various studies have shown that different alterations in the WNT signaling pathway can be concentrated in different EO-CRC subgroups, depending on molecular characteristics such as MSI status and mutation burden [6,18,21,27].

The most common BRAF mutation is BRAFV600E. Literature shows that BRAFV600E mutations are found in approximately 5-10% of patients with colorectal carcinoma [28]. In our study, BRAF mutation was found to be significantly more frequent in LO-CRC than EO-CRC, which is largely consistent with the literature. It is known that BRAF V600E mutations are more common in older patients with colorectal carcinoma [3,7,19,25]. Several studies comparing early-onset and late-onset CRC have also reported that BRAF mutations are less frequent in younger patients [6,7,17-19,29].

In our study, RNF43 mutations were also detected significantly more frequently in the LO-CRC group. RNF43 is one of the negative regulators of WNT signaling and can undergo recurrent mutations in colorectal tumors. Yan et al. showed a close association between RNF43 mutations and serrated neoplasia and BRAF mutations [30].

In this study we also found that mutations in AXIN2, PLCG2, AMER1, MSH3, MGA, SOX9, and SPEN were significantly more frequent in the LO-CRC group. While the clinical significance of age-related differences in these genes is not yet clear, some are associated with known molecular pathways in colorectal cancer pathogenesis [27,31]. AXIN2 and AMER1 are involved in the regulation of WNT/β-catenin signaling, while SOX9 plays a role in intestinal epithelial differentiation and stem cell biology. AMER1 was previously identified as one of the recurrently mutated genes in colorectal cancer through exome sequencing studies [32-34]. MSH3 is involved in the DNA mismatch repair mechanism, and changes in this gene may be associated with certain genomic instability phenotypes [35]. SPEN is a gene involved in transcriptional regulation, and recent data suggest that SPEN mutations may be a potential biomarker associated with immunotherapy response in colorectal cancer [36]. Large-scale studies are needed to more clearly define the effects of these genes in patients with colorectal cancer.

In our study, male patients constituted the majority in both the EO-CRC and LO-CRC groups; however, the proportion of males was significantly higher in the LO-CRC group. Previous studies have reported conflicting results regarding gender distribution in EO-CRC; some cohorts show a male predominance, while others

indicate a higher proportion of women or a correlation between female gender and younger age at diagnosis. In the study by Park et al., 52.5% of cases in the 20-49 age range were male [8]. Siddique et al. reported that male gender was associated with an increased risk of early-onset colorectal cancer [37]. Abdel-Rahman et al. showed that colorectal cancer diagnosis at a younger age was associated with female gender [38].

In our study, the median mutational count was significantly lower in patients with EO-CRC than in those with LO-CRC. Similarly, the median fraction of genome altered was slightly lower in the EO-CRC group than in the LO-CRC group. Although these differences were statistically significant, their absolute magnitudes were modest, and their clinical relevance remains uncertain.

This study has some limitations. Primarily, due to its retrospective design, there are gaps in clinical features such as family history, tumor stage, lymph node involvement, metastasis status, patient survival time, and treatment modalities. Furthermore, genomic assessments such as tumor mutational burden, mRNA sequencing, and miRNA analysis are not available for these patients. Additionally, the inclusion of both primary and metastatic tumor samples may have introduced heterogeneity into the genomic comparisons, as the mutational profiles of metastatic tumors may differ from those of primary tumors due to tumor evolution and treatment-related selective pressures. Furthermore, molecular subgroups based on microsatellite instability (MSI) or microsatellite stability (MSS) could not be adequately controlled for in the present analysis. As MSI status is associated with distinct genomic profiles and may influence the frequency of several mutations, the lack of stratification according to MSI/MSS status may have affected the observed differences between the EO-CRC and LO-CRC groups.

In conclusion, our study demonstrates measurable differences in somatic genomic profiles between EO-CRC and LO-CRC. The higher frequency of TP53, APC, and TCF7L2 mutations in EO-CRC, and conversely, the enrichment of BRAF, RNF43, and various other gene alterations in LO-CRC, suggest that the relative contribution of molecular pathways involved in colorectal carcinogenesis may differ according to age. However, EO-CRC should not be considered a completely separate molecular disease based solely on age. These molecular differences are important in terms of their potential to guide future screening, prognosis, or personalized treatments.

Author contributions

Conception and design: B.Ö., O.K., T.K.; Data acquisition: BÖ; Data analysis: B.Ö.; Data interpretation: B.Ö.; Drafting of the manuscript: B.Ö.; Critical revision of the manuscript: H.A.Y., O.K., T.K. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was conducted using de-identified clinical and genomic information available through the American Association for Cancer Research (AACR) Project GENIE registry. Protection of participant privacy and confidentiality is governed by the ethical and regulatory procedures of the GENIE consortium and the Institutional Review Boards (IRBs) of the contributing institutions. Depending on local institutional policies, submission of data to the registry is permitted through informed consent for sharing de-identified molecular information, an IRB-approved waiver of consent, or institutionally approved GENIE research protocols. All patient-level information is de-identified before being incorporated into the registry in accordance with the HIPAA Safe Harbor standards. Because the current analysis relied solely on previously collected and de-identified registry data, investigators had no direct interaction with participants and no additional specimens or biological materials were obtained.

Data availability statement

Data sharing is not applicable to this article as no new datasets were generated or analyzed during this study.

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

Acknowledgments

The authors gratefully acknowledge the American Association for Cancer Research (AACR) for the development and continued support of the AACR Project GENIE registry. We also thank all participating institutions and consortium members for contributing genomic and clinical data that enabled this research. The findings, interpretations, and conclusions reported herein are the responsibility of the authors and do not necessarily reflect those of AACR Project GENIE or its participating institutions.

References

[1] Sung H, Filho AM, Laversanne M, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 34 cancers in 186 countries. CA Cancer J Clin 2026;76(4):e70090. **[Crossref]**

[2] Morgan E, Arnold M, Gini A, et al. Global burden of colorectal cancer in 2020 and 2040: incidence and mortality estimates from GLOBOCAN. Gut 2023;72(2):338-44. **[Crossref]**

[3] Saraiva MR, Rosa I, Claro I. Early-onset colorectal cancer: a review of current knowledge. World J Gastroenterol 2023;29(8):1289-303. **[Crossref]**

[4] Turk A, Mondaca S, Nervi B, Morris AD, Finer Z, Holowatyj AN. Early-onset colorectal cancer: from genetic discovery to clinical innovation. Am Soc Clin Oncol Educ Book 2025;45(3):e473618. **[Crossref]**

[5] Sinicrope FA. Increasing incidence of early-onset colorectal cancer. N Engl J Med 2022;386(16):1547-58. **[Crossref]**

[6] Li J, Pan Y, Guo F, et al. Patterns in genomic mutations among patients with early-onset colorectal cancer: an international, multicohort, observational study. Lancet Oncol 2025;26(8):1055-66. **[Crossref]**

[7] Cercek A, Chatila WK, Yaeger R, et al. A comprehensive comparison of early-onset and average-onset colorectal cancers. J Natl Cancer Inst 2021;113(12):1683-92. **[Crossref]**

[8] Park L, O'Connell K, Herzog K, et al. Clinical features of young onset colorectal cancer patients from a large cohort at a single cancer center. Int J Colorectal Dis 2022;37(12):2511-6. **[Crossref]**

[9] Peltomäki P, Nyström M, Mecklin JP, Seppälä TT. Lynch syndrome genetics and clinical implications. Gastroenterology 2023;164(5):783-99. **[Crossref]**

[10] Aelvoet AS, Buttitta F, Ricciardiello L, Dekker E. Management of familial adenomatous polyposis and MUTYH-associated polyposis; new insights. Best Pract Res Clin Gastroenterol 2022;58-59:101793. **[Crossref]**

[11] Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov 2012;2(5):401-4. **[Crossref]**

[12] Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. Sci Signal 2013;6(269):pl1. **[Crossref]**

[13] AACR Project GENIE Consortium. AACR Project GENIE: powering precision medicine through an international consortium. Cancer Discov 2017;7(8):818-31. **[Crossref]**

[14] AACR Project GENIE. AACR Project GENIE Data. 2026. Available at: https://aacrprojectgenie.org/data/

[15] de Bruijn I, Kundra R, Mastrogiacomo B, et al. Analysis and visualization of longitudinal genomic and clinical data from the AACR project GENIE biopharma collaborative in cBioPortal. Cancer Res 2023;83(23):3861-7. **[Crossref]**

[16] Yan S, Zhan F, He Y, Zhu Y, Ma Z. p53 in colorectal cancer: from a master player to a privileged therapy target. J Transl Med 2025;23(1):684. **[Crossref]**

[17] Tang J, Peng W, Tian C, et al. Molecular characteristics of early-onset compared with late-onset colorectal cancer: a case controlled study. Int J Surg 2024;110(8):4559-70. **[Crossref]**

[18] Storandt MH, Shi Q, Eng C, et al. Genomic landscapes of early-onset versus average-onset colorectal cancer populations. Cancers (Basel) 2025;17(5):836. **[Crossref]**

[19] Lieu CH, Golemis EA, Serebriiskii IG, et al. Comprehensive genomic landscapes in early and later onset colorectal cancer. Clin Cancer Res 2019;25(19):5852-8. **[Crossref]**

[20] Lawler T, Parlato L, Warren Andersen S. The histological and molecular characteristics of early-onset colorectal cancer: a systematic review and meta-analysis. Front Oncol 2024;14:1349572. **[Crossref]**

[21] Pretta A, Nasca V, Marmorino F, et al. Early-onset colorectal cancer patients exhibit a distinct molecular fingerprint: insights from a large-scale NGS study of 1209 patients. ESMO Open 2025;10(9):105756. **[Crossref]**

[22] Markowitz SD, Bertagnolli MM. Molecular origins of cancer: molecular basis of colorectal cancer. N Engl J Med 2009;361(25):2449-60. **[Crossref]**

[23] Preisler L, Habib A, Shapira G, et al. Heterozygous APC germline mutations impart predisposition to colorectal cancer. Sci Rep 2021;11(1):5113. **[Crossref]**

[24] Alshenaifi JY, Vetere G, Maddalena G, et al. Mutational and co-mutational landscape of early onset colorectal cancer. Biomarkers 2025;30(1):64-76. **[Crossref]**

[25] Cancer Genome Atlas Network. Comprehensive molecular characterization of human colon and rectal cancer. Nature 2012;487(7407):330-7. **[Crossref]**

[26] Wenzel J, Rose K, Haghighi EB, et al. Loss of the nuclear Wnt pathway effector TCF7L2 promotes migration and invasion of human colorectal cancer cells. Oncogene 2020;39(19):3893-909. **[Crossref]**

[27] Holowatyj AN, Wen W, Gibbs T, et al. Racial/ethnic and sex differences in somatic cancer gene mutations among patients with early-onset colorectal cancer. Cancer Discov 2023;13(3):570-9. **[Crossref]**

[28] Lai E, Pretta A, Impera V, et al. BRAF-mutant colorectal cancer, a different breed evolving. Expert Rev Mol Diagn 2018;18(6):499-512. **[Crossref]**

[29] Willauer AN, Liu Y, Pereira AAL, et al. Clinical and molecular characterization of early-onset colorectal cancer. Cancer 2019;125(12):2002-10. **[Crossref]**

[30] Yan HHN, Lai JCW, Ho SL, et al. RNF43 germline and somatic mutation in serrated neoplasia pathway and its association with BRAF mutation. Gut 2017;66(9):1645-56. **[Crossref]**

[31] Cavestro GM, Mannucci A, Balaguer F, et al. Delphi Initiative for Early-Onset Colorectal Cancer (DIRECt) international management guidelines. Clin Gastroenterol Hepatol 2023;21(3):581-603.e33. **[Crossref]**

[32] Sanz-Pamplona R, Lopez-Doriga A, Paré-Brunet L, et al. Exome sequencing reveals AMER1 as a frequently mutated gene in colorectal cancer. Clin Cancer Res 2015;21(20):4709-18. **[Crossref]**

[33] Rebuzzi F, Ulivi P, Tedaldi G. Genetic predisposition to colorectal cancer: how many and which genes to test? Int J Mol Sci 2023;24(3):2137. **[Crossref]**

[34] Lizárraga-Verdugo E, Ruiz-García E, López-Camarillo C, et al. Cell survival is regulated via SOX9/BCL2L1 axis in HCT-116 colorectal cancer cell line. J Oncol 2020;2020:5701527. **[Crossref]**

[35] Laycock A, Kang A, Ang S, Texler M, Bentel J. Lack of correlation between MSH3 immunohistochemistry and microsatellite analysis for the detection of elevated microsatellite alterations at selected tetranucleotide repeats (EMAST) in colorectal cancers. Hum Pathol 2021;118:9-17. **[Crossref]**

[36] Dong Y, Ye S, Li H, Li J, Liu R, Zhu Y. The role of SPEN mutations as predictive biomarkers for immunotherapy response in colorectal cancer: insights from a retrospective cohort analysis. J Pers Med 2024;14(2):131. **[Crossref]**

[37] Siddique S, Wang R, Berardi D, et al. Demographic, birth, parental characteristics, and the risk of early-onset colorectal cancer: a population-based nested case-control study in California. Cancer 2026;132(13):e70458. **[Crossref]**

[38] Abdel-Rahman O, Karachiwala H, Koski S. Patterns of colorectal cancer diagnosis among younger adults in a real-world, population-based cohort. Future Oncol 2022;18(1):47-54. **[Crossref]**

Acta Medica 2026; 57(3): 266-272 DOI: 10.32552/actamedica.2026.1244

acta medica

RESEARCH ARTICLE

Outcomes of oral provocation tests with alternative iron preparations in patients with suspected iron hypersensitivity

Özge Öztürk Aktaş[1], Ebru Damadoğlu[1], Gül Karakaya[1], Ali Fuat Kalyoncu[1]

[1]Division of Allergy and Clinical Immunology, Department of Chest Diseases, School of Medicine, Hacettepe University, Ankara, Türkiye

Abstract

Objective: Although hypersensitivity reactions to oral iron preparations are rare, they are clinically challenging as they can limit treatment options for iron deficiency anemia. Data on cross-reactivity between different oral iron formulations are limited, and guidance on safe switching strategies is insufficient. This study aimed to define the clinical characteristics of patients with suspected oral iron hypersensitivity and to evaluate the tolerance of alternative oral iron preparations using supervised oral provocation testing.

Materials and Methods: We retrospectively analyzed 18 adult patients who presented to a tertiary allergy clinic with a history of hypersensitivity to an oral iron preparation between June 2011 and June 2025. Alternative oral iron formulations were selected based on differences in iron valence ferrous (Fe^{2+}) and ferric (Fe^{3+}) and ligand or complex structure. Oral provocation tests were performed using graded protocols.

Results: The cohort included 17 females and 1 male with a mean age of 42.7 ± 9.5 years. Most reactions were attributed to ferrous (Fe^{2+}) salts (77.8%), whereas ferric (Fe^{3+}) salts accounted for 22.2% of the culprit oral iron formulations. The most frequent symptom was skin symptoms (77.8%), followed by angioedema or mucosal involvement (38.9%), systemic or vasovagal-like symptoms (27.8%), and gastrointestinal complaints (11.1%). Three patients experienced clinically anaphylactic reactions. Alternative iron formulations were selected by altering the iron oxidation status and/or ligand structure. All oral provocation tests with alternative preparations, including extended provocation protocols, were negative (100%), indicating excellent tolerance and a low probability of clinically significant cross-reaction.

Conclusion: In patients suspected of having oral iron hypersensitivity, switching to an alternative oral iron preparation with a different oxidation state and/or ligand by performing provocation tests appears to be a safe and effective strategy. These findings support that hypersensitivity reactions are more closely related to formulation-specific components than to iron itself, and offer a practical approach to maintaining oral iron therapy.

Keywords: drug hypersensitivity, iron compounds, drug provocation tests

Introduction

Iron deficiency anemia is a prevalent nutritional condition globally, impacting up to 25% of the population, particularly among children, women of reproductive age, and those in low and middle-income nations [1]. Allergic reactions to oral iron are uncommon; most side effects are gastrointestinal rather than true hypersensitivity. Intravenous iron is more likely to cause severe allergic reactions, such as anaphylaxis, than oral iron [2]. Complement activation-related pseudo-allergy induced by iron nanoparticles is likely a more

Corresponding author: Özge Öztürk Aktaş • **Email:** ozge.ozturk@hacettepe.edu.tr

Received: January 8, 2026 **Accepted:** August 10, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

prevalent pathogenic mechanism in acute reactions to contemporary intravenous iron formulations than an immunological IgE-mediated response [3]. According to ENDA/EAACI standards, individuals with suspected drug allergies should have a thorough clinical history assessment, subsequently followed by drug provocation testing with the suspected or an alternate medication, when suitable [4]. In patients with a history of severe acute reactions, drug provocation testing with the culprit drug is generally avoided; however, supervised oral provocation with an alternative formulation may represent a safe and effective approach. Patients experiencing hypersensitivity reactions to iron preparations pose a considerable clinical problem, since these reactions may restrict viable therapy alternatives for iron deficiency anemia. Furthermore, information regarding cross-reactivity among various oral iron salts is limited, complicating the prediction of tolerance to alternate iron formulations and hindering clinical decision-making. The aim of this study was to characterize the clinical patterns of reactions attributed to suspected oral iron allergy and to evaluate the tolerability of alternative oral iron preparations using oral provocation tests.

Materials and Methods

Study design

The characteristics of 18 patients who attended our Adult Allergy Clinic due to allergies to oral iron preparations from June 2011 to June 2025, along with the outcomes of provocation tests conducted with other oral iron medications, were analyzed.

Patients were included in the study if they had a clinical history suggestive of a hypersensitivity reaction to an oral iron formulation. The research was conducted at a tertiary adult allergy clinic. After a detailed clinical history, patients were evaluated per standard clinical procedures, and when appropriate, oral provocation testing with a different oral iron formulation was conducted to evaluate tolerance. The choice of an alternate iron salt was predicated on variations in iron valency and conjugate configuration. The study protocol was approved by the local ethics committee. (Approval No: SBA25/1024; Date: 09 Dec 2025)

Data collection and oral provocation tests

Demographic characteristics including age and sex, the suspected culprit oral iron preparation, and detailed descriptions of clinical reaction patterns were recorded. Clinical reactions were classified according to their type and timing, as immediate or delayed. The following symptoms were systematically evaluated during the hypersensitivity reaction assessment: cutaneous manifestations, including flushing, pruritus, urticaria, and angioedema; cardiovascular symptoms such as chest pain, tachycardia, presyncope, syncope, and hypotension; respiratory features including nasopharyngeal symptoms, dyspnea, wheezing, and oxygen desaturation; throat tightness; gastrointestinal complaints including nausea, vomiting, diarrhea, and abdominal pain; and atypical manifestations such as fever or chills, back and neck pain, and sensory or motor disturbances.

The drug provocation test (DPT) was considered the reference standard for confirming or excluding hypersensitivity to iron preparations and were performed in accordance with the European Network for Drug Allergy (ENDA) guidelines and the recommendations of the EAACI Drug Allergy Interest Group [4]. In cases where the clinical association between the hypersensitivity reaction and the culprit iron formulation was strong, particularly when the initial reaction had been severe, testing was conducted using an alternative iron preparation rather than the suspected agent. Alternative oral iron preparations were selected based on differences in iron valency ferrous (Fe^{2+}) versus ferric (Fe^{3+}) formulations and conjugate or excipient structure, aiming to minimize potential cross-reactivity.

Provocation tests were conducted using graded oral challenge protocol, depending on the clinical history and severity of the reported reaction. DPTs were carried out using a graded dosing protocol, starting at 1/100 of the therapeutic dose, followed by stepwise increases to 1/10, 1/2, and finally the full dose at 30–60 minute intervals. All tests were performed under medical supervision with an appropriate observation period following administration. After administration of the cumulative target dose, patients were observed for at least four hours.

A test was considered positive if objective clinical signs consistent with hypersensitivity occurred during or after the challenge. In the absence of any clinical symptoms, the test result was defined as negative.

Statistical analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 27.0 (IBM Corp., Armonk, NY, USA). Continuous variables were summarized as mean ± standard deviation (SD) for normally distributed data and as median with interquartile range (IQR) for non-normally distributed data, as appropriate. Categorical variables were summarized as frequencies and percentages. Given the descriptive nature of this retrospective study and the limited sample size, the analyses were primarily descriptive.

Results

Patient characteristics and initial culprit iron preparations

The study population consisted of 18 patients with a mean age of 42.7 ± 9.5 years. The majority of patients were female (n = 17, 94.4%), while only one patient (5.6%) was male (Table 1).

Among the 18 patients included in the analysis, hypersensitivity reactions were most frequently associated with oral ferrous iron preparations, accounting for 14 cases (77.8%). These reactions were mainly attributed to ferrous glycine sulfate, ferrous fumarate, and ferrous sulfate formulations. In contrast, ferric oral iron preparations were implicated in 4 patients (22.2%), predominantly involving ferric hydroxide polymaltose and ferric protein succinylate. No patient experienced a reaction to intravenous iron in this updated cohort (Figure 1).

Table 1. Baseline characteristics and reaction patterns of patients with suspected oral iron hypersensitivity.

Variable	Value
Total patients, n	18
Female, n (%)	17 (94.4%)
Male, n (%)	1 (5.6%)
Age, mean ± SD (years)	42.7 ± 9.5
Culprit iron formulation	
Ferrous (Fe^{2+}) salts, n (%)	14 (77.8%)
Ferric (Fe^{3+}) salts, n (%)	4 (22.2%)
Most frequent culprit formulation	Ferrous glycine sulfate
Clinical phenotype of reactions*	
Cutaneous (pruritus, urticaria, rash), n (%)	14 (77.8%)
Angioedema / mucosal involvement, n (%)	7 (38.9%)
Vasovagal or systemic symptoms, n (%)	5 (27.8%)
Gastrointestinal symptoms, n (%)	2 (11.1%)
Anaphylaxis, n (%)	3 (16.7%)
Timing of hypersensitivity reaction	
Early onset (≤ 6 hours), n (%)	9 (50.0%)
Delayed onset (> 6 hours), n (%)	9 (50.0%)

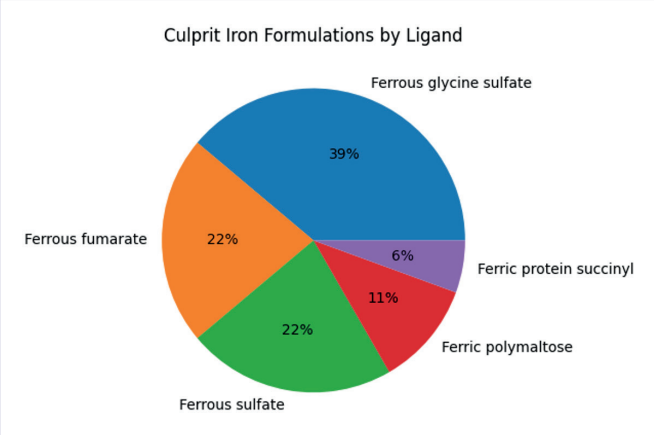


Figure 1. Distribution of culprit oral iron formulations according to ligand and complex structure.

Clinical characteristics of hypersensitivity reactions

The most common clinical presentation was cutaneous involvement, including pruritus, erythema, urticaria, or maculopapular eruptions, observed in 14 patients (77.8%). Angioedema or mucosal involvement, such as lip, oral cavity or throat swelling, hoarseness, or dyspnea, was documented in 7 patients (38.9%). Systemic or vasovagal-like reactions, including dizziness, palpitations, sweating, syncope, or transient loss of consciousness, occurred in 5 patients (27.8%), while gastrointestinal symptoms (nausea and/or diarrhea) accompanied the reaction in 2 patients (11.1%). Several patients exhibited overlapping symptom patterns. Three patients' reactions were compatible with anaphylaxis.

This cohort exhibited two patterns of hypersensitive reactions to oral iron. Nine out of 18 patients (50%) had symptoms within the initial 6 hours post-drug administration and were categorized as experiencing early responses. These responses generally manifested following the initial or subsequent dose and included dizziness, nausea, pruritus, angioedema, respiratory problems, or syncope. Nine patients (50%) showed delayed reactions, with symptom onset following repeated dosage or several days of exposure, predominantly between days 3 and 10. Delayed reactions were mainly marked by cutaneous symptoms including itching, urticaria, rashes, and localized edema.

Alternative iron preparations and oral provocation tests

We selected alternative oral iron formulations based on both the oxidation state of iron (Fe^{2+} vs. Fe^{3+}) and the ligand/side-chain structure of the complex. Ferric (Fe^{3+}) preparations constituted the majority of alternative oral agents used in provocation testing (n=10/18, 55.6%). Among these, ferric protein succinylate was administered most frequently (n=5, 27.8%), followed by ferric hydroxide polymaltose complexes (n=4, 22.2%) and ferric maltol in one case (n=1, 5.6%). Ferrous formulations were selected in 8 patients (44.4%), including ferrous fumarate (n=3, 16.7%), ferrous glycinate sulfate (n=2, 11.1%), ferrous gluconate (n=2, 11.1%), and ferrous sulfate in one patient (n=1, 5.6%) (Figure 2).

Taken together, these results demonstrate that alternative compounds with distinct oxidation states and ligand complexes were intentionally selected to minimize the risk of recurrent hypersensitivity reactions. All oral provocation tests were negative (100%), including extended protocols where applicable, demonstrating good tolerability of alternative iron formulations and supporting a low likelihood of clinically relevant cross-reactivity between different oral iron compounds.

Discussion

We aimed to demonstrate the outcome of using alternative oral iron preparations in patients with suspected hypersensitivity reactions and to provide a practical, clinically applicable approach for continued iron replacement when treatment remains indicated. Accordingly, our focus was on the tolerability of switching strategies supported by supervised oral provocation testing rather than on identifying underlying predisposing risk factors.

Hypersensitivity reactions to iron have been described predominantly with intravenous formulations, with broad clinical spectrum from mild, self-limited symptoms to severe systemic reactions. Adverse events may be influenced by the infusion rate and can present with dyspnea, chest or back pain, musculoskeletal discomfort,

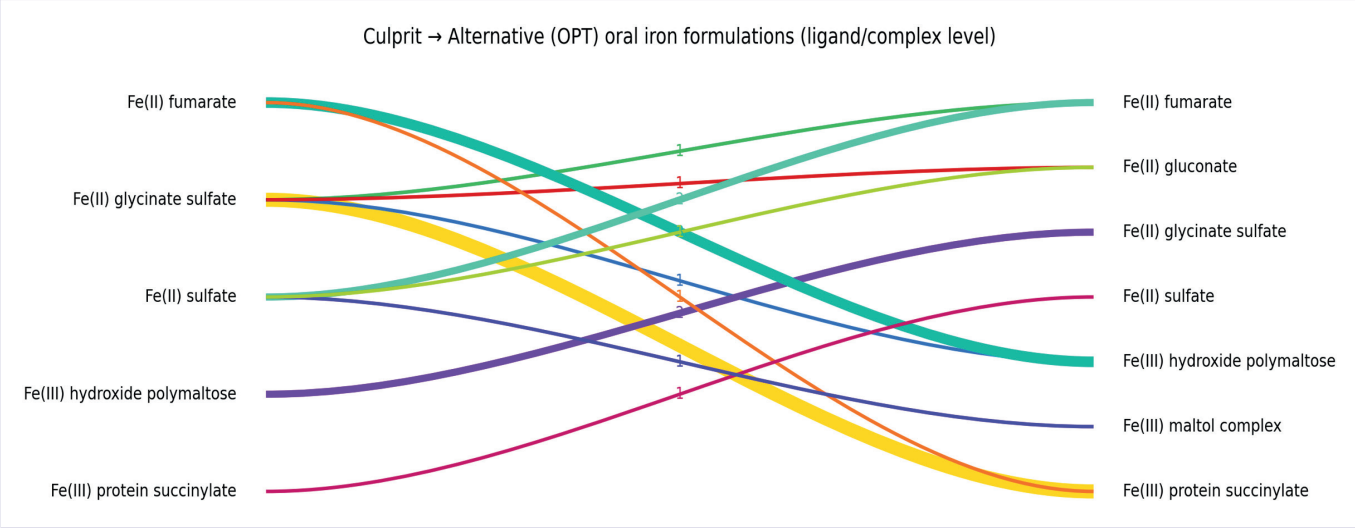


Figure 2. Flow of patients from culprit to alternative oral iron formulations during oral provocation testing, stratified by iron valence and ligand/complex.

hypotension, tachycardia, nausea, or pruritus [5]. Mild reactions often resolve spontaneously or after slowing the infusion and may not recur on rechallenge, whereas manifestations such as angioedema, urticaria, or severe systemic reactions including anaphylaxis warrant careful evaluation. Delayed presentations, characterized by hypotension, myalgias, malaise, or vomiting have also been reported [6-8]. Atopy and various medication allergies, especially past iron hypersensitivity reactions, are risk factors for severe hypersensitivity reactions [9].

Only a limited number of cases have been published describing hypersensitivity reactions to oral iron preparations. Reported phenotypes include maculopapular eruptions, pustular eruptions, vasculitis, and photodermatitis [10-13]. Urticaria with or without angioedema has also been described. In rare instances, anaphylaxis has been reported [14].

Although oral preparations are generally better tolerated and less frequently associated with hypersensitivity, any suspected reaction to oral iron should nevertheless be evaluated and managed carefully [8]. In line with current international guidelines, we avoided re-exposure to the culprit iron preparation in patients with a history of severe immediate reactions, particularly anaphylaxis, and preferentially selected alternative oral iron formulations whenever possible [4,15].

In line with prior findings on iron-induced hypersensitivity, skin manifestations were the primary phenotype in our cohort, with pruritus, erythema, urticaria, and maculopapular eruptions occurring most frequently, while angioedema and mucosal symptoms (swelling of the lips, oral cavity, or throat, hoarseness, dyspnea) were also an important feature. A small number of patients experienced gastrointestinal problems, such as nausea or diarrhea, during the reaction, indicating that non-cutaneous symptoms may accompany skin manifestations in oral iron hypersensitivity reactions. A limited number of cases were clinically consistent with anaphylaxis; however, none of these instances underwent confirmatory testing with the same offending agent. This was mainly due to the practical accessibility of various alternative oral iron formulations, limited patient consent to undergo re-exposure to the suspected agent, and the absence of standardized diagnostic testing for most oral iron products in regular practice. Our findings indicate that oral iron hypersensitivity may present with both immediate and delayed reaction patterns. While immediate reactions were more often characterized by systemic manifestations such as angioedema, respiratory symptoms, or syncope, delayed reactions predominantly presented with cutaneous manifestations, highlighting the heterogeneous clinical spectrum of oral iron hypersensitivity. This distribution suggests that both acute and delayed hypersensitivity reactions are prevalent in patients with suspected oral iron allergy, underscoring the necessity for tailored monitoring measures during assessment and drug provocation testing.

In most published reports, either ferrous or ferric iron preparations have been identified as the suspected triggers. Ferric salts are considered the most chemically stable form of iron [16]. Greater stability may reduce binding to macromolecules, which could translate into lower allergenic potential. Ferric formulations are also generally less absorbable than ferrous salts. Reduced absorption may limit systemic exposure and thereby decrease the likelihood of clinically apparent reactions. This may partly explain why allergic reactions appear to be reported less often with ferric than with ferrous preparations.

Cross-reactivity has been documented among several ferrous salts, including ferrous sulfate, ferrous ascorbate, and ferrous lactate [3]. A case report described a patient who developed hypersensitivity to ferrous glycine sulfate and ferrous fumarate, yet tolerated ferric polymaltose complex, suggesting that differences in oxidation state and the associated ligand may be clinically meaningful when selecting an alternative oral iron agent [17]. These findings help explain why changing to an iron preparation with a different oxidation state or ligand can be clinically useful in patients with suspected oral iron hypersensitivity. They also support a stepwise approach to selecting an alternative formulation before considering desensitization. Similarly, a pediatric case series demonstrated that four children with hypersensitivity to oral iron preparations successfully tolerated alternative oral iron formulations following supervised oral provocation testing, further supporting the selection of an alternative oral iron preparation in patients with suspected oral iron hypersensitivity [18].

A potential nonspecific immunological mechanism for hypersensitivity reactions to nondextran iron products is complement activation-related pseudoallergy (CARPA). In CARPA, complement activation produces anaphylatoxins, notably C3a and C5a, which can activate effector cells such mast cells and basophils, resulting in clinical manifestations similar to acute hypersensitivity [19,20].

In our study, oral provocation testing was successful in all patients when an alternative preparation was selected by changing the iron valence and/or the accompanying compound (ligand/complex). This finding supports the concept that, in suspected iron allergy reactions may be more closely related to the formulation components (i.e., the ligand/complex or excipients) than to elemental iron itself. In our cohort, the choice of an alternative iron preparation considered both the oxidation state (ferrous vs. ferric) and the ligand/complex (e.g., polymaltose, protein succinylate, maltol, gluconate, glycinate) to reduce the likelihood of clinically relevant cross-reactivity. Our findings indicate that alternative oral iron formulations can frequently be safely administered.

Author contributions

Conception and design: Ö.Ö.A., A.F.K.; Conception: E.D., G.K.; Data acquisition: Ö.Ö.A., E.D., G.K.; Data analysis: Ö.Ö.A., E.D., G.K., A.F.K.; Data interpretation: Ö.Ö.A., E.D., G.K., A.F.K.; Drafting of the manuscript: Ö.Ö.A., E.D., G.K., A.F.K.; Critical revision of the manuscript: Ö.Ö.A., E.D., G.K., A.F.K. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Hacettepe University ethics committee (Date: December 9, 2025, Decision/ Protocol No: SBA25/1024). Informed consent was obtained from all participants involved in this study.

Data availability statement

The data supporting the findings of this study are not publicly available due to containing information that could compromise the privacy of research participants.

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

References

[1] Kolarš B, Mijatović Jovin V, Živanović N, et al. Iron deficiency and iron deficiency anemia: a comprehensive overview of established and emerging concepts. Pharmaceuticals (Basel) 2025;18(8):1104. **[Crossref]**

[2] Pantopoulos K. Oral iron supplementation: new formulations, old questions. Haematologica 2024;109(9):2790-801. **[Crossref]**

[3] Rampton D, Folkersen J, Fishbane S, et al. Hypersensitivity reactions to intravenous iron: guidance for risk minimization and management. Haematologica 2014;99(11):1671-6. **[Crossref]**

[4] Barbaud A, Garvey LH, Torres M, et al. EAACI/ENDA position paper on drug provocation testing. Allergy 2024;79(3):565-79. **[Crossref]**

[5] Gómez-Ramírez S, Shander A, Spahn DR, et al. Prevention and management of acute reactions to intravenous iron in surgical patients. Blood Transfus 2019;17(2):137-45. **[Crossref]**

[6] Barton JC, Barton EH, Bertoli LF, Gothard CH, Sherrer JS. Intravenous iron dextran therapy in patients with iron deficiency and normal renal function who failed to respond to or did not tolerate oral iron supplementation. Am J Med 2000;109(1):27-32. **[Crossref]**

[7] Fishbane S, Ungureanu VD, Maesaka JK, Kaupke CJ, Lim V, Wish J. The safety of intravenous iron dextran in hemodialysis patients. Am J Kidney Dis 1996;28(4):529-34. **[Crossref]**

[8] Wysowski DK, Swartz L, Borders-Hemphill BV, Goulding MR, Dormitzer C. Use of parenteral iron products and serious anaphylactic-type reactions. Am J Hematol 2010;85(9):650-4. **[Crossref]**

[9] Szebeni J, Fishbane S, Hedenus M, et al. Hypersensitivity to intravenous iron: classification, terminology, mechanisms and management. Br J Pharmacol 2015;172(21):5025-36. **[Crossref]**

[10] Ortega N, Castillo R, Blanco C, Alvarez M, Carrillo T. Oral iron cutaneous adverse reaction and successful desensitization. Ann Allergy Asthma Immunol 2000;84(1):43-5. **[Crossref]**

[11] Niemi KM, Kangas K. Non-thrombocytopenic purpuras. Acta Derm Venereol 1978;58:337-42. **[Crossref]**

[12] Ito A, Nomura K, Hashimoto I. Pustular drug eruption induced by ferrous fumarate. Dermatology 1996;192(3):294-5.

[13] Rodríguez-Jiménez B, Domínguez-Ortega J, Nuñez-Acevedo B, Cava-Sumner B, Kindelan-Recarte C, Montojo-Guillén C. Rapid iron desensitization after generalized urticaria and facial angioedema. J Investig Allergol Clin Immunol 2014;24(1):69-71.

[14] de Barrio M, Fuentes V, Tornero P, Sánchez I, Zubeldia J, Herrero T. Anaphylaxis to oral iron salts. desensitization protocol for tolerance induction. J Investig Allergol Clin Immunol 2008;18(4):305-8.

[15] Golden DBK, Wang J, Waserman S, et al. Anaphylaxis: a 2023 practice parameter update. Ann Allergy Asthma Immunol 2024;132(2):124-76. **[Crossref]**

[16] Cancelo-Hidalgo MJ, Castelo-Branco C, Palacios S, et al. Tolerability of different oral iron supplements: a systematic review. Curr Med Res Opin 2013;29(4):291-303. **[Crossref]**

[17] Yilmaz O, Erkasar Citak F. A case report with hypersensitivity to ferrous but tolerance of ferric iron salts. J Allergy Clin Immunol Pract 2016;4(2):343-4.e1. **[Crossref]**

[18] Özdemir Ö, Büyükavcı M. Development of hypersensitivity reactions after using different oral iron preparations. Istanbul Med J 2018;19:173-5. **[Crossref]**

[19] Macdougall IC, Vernon K. Complement activation-related pseudo-allergy: a fresh look at hypersensitivity reactions to intravenous iron. Am J Nephrol 2017;45(1):60-2. **[Crossref]**

[20] Szebeni J, Storm G. Complement activation as a bioequivalence issue relevant to the development of generic liposomes and other nanoparticulate drugs. Biochem Biophys Res Commun 2015;468(3):490-7. **[Crossref]**

Acta Medica 2026; 57(3): 273-282 DOI: 10.32552/actamedica.2026.1330

acta medica

RESEARCH ARTICLE

Intensive care unit admissions in adult cystic fibrosis patients: A national cohort study from Türkiye

Mehmet Yıldırım[1], Emin Gemcioğlu[2], Rıfat Bozkuş[2], Şuayip Birinci[3], Arzu Topeli[1], Naim Ata[4]

[1]Division of Intensive Care, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
[2]Department of Internal Medicine, Faculty of Medicine, Health Sciences University, Ankara, Türkiye
[3]Deputy Minister, Republic of Türkiye Ministry of Health, Ankara, Türkiye
[4]General Directorate of Information Systems, Ministry of Health, Ankara, Türkiye

Abstract

Objective: Improvements in survival have led to a growing adult population with cystic fibrosis (CF), increasing the relevance of intensive care unit (ICU) management in this group. However, nationwide data on ICU utilization and outcomes in CF remain limited.

Materials and Methods: This nationwide retrospective cohort study included all adults (≥18 years) with CF admitted to ICUs in Türkiye between 2018 and 2023. Data were obtained from the Turkish Ministry of Health National Electronic Database, covering all ICU levels and hospital types. Demographic characteristics, admission diagnoses, comorbidities, ICU interventions, APACHE II scores, and outcomes were analyzed. The primary outcome was ICU mortality.

Results: A total of 126 adults with CF were admitted to ICUs during the study period. Most admissions occurred in Level 3 ICUs (62.7%) and mixed ICUs (38.8%). More than half were concentrated in the Marmara region (50.8%). Median age was 30 years [23–41], and overall ICU mortality was 32.5%. Non-survivors were older and had higher APACHE II scores than survivors (21 [12–28] vs. 7 [3–13], p<0.01). Sepsis was more frequent among non-survivors (31.7% vs. 15.3%, p=0.03), and cardiovascular comorbidities were associated with mortality. Non-invasive mechanical ventilation was commonly used but was not associated with ICU mortality.

Conclusions: In this nationwide cohort, ICU mortality among adults with CF was comparable to contemporary international reports and was primarily associated with acute illness severity and infectious complications. These findings provide the first national overview of ICU care in CF in Türkiye and highlight the importance of severity-based risk stratification in critically ill adults with CF.

Keywords: adult, acute respiratory failure, critical care outcomes, critical illness, cystic fibrosis, mechanical ventilation, mortality

Introduction

Cystic fibrosis (CF) is a multisystem genetic disorder characterized by chronic pulmonary disease, recurrent infections, and progressive organ dysfunction [1]. Advances in respiratory care, antimicrobial therapy, and multidisciplinary follow-up have substantially improved survival, leading to an increasing number of adult patients living with CF [2,3]. This demographic shift is particularly evident in established national registries; in the United States, adults accounted for 61.6% of individuals included in the Cystic Fibrosis Foundation Patient Registry in 2024 [4]. A similar transition is emerging in Türkiye. According to the Cystic Fibrosis

Corresponding author: Mehmet Yıldırım • **Email:** dryildirimmehmet@hotmail.com

Received: June 25, 2026 **Accepted:** August 26, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Registry of Türkiye, 2,258 patients were registered in 2023, of whom 336 (14.9%) were aged ≥18 years, compared with only nine registered adult patients in 2016 [5]. As the number of adults living with CF increases, understanding the patterns and outcomes of critical illness in this population becomes increasingly relevant. Advances in respiratory care, antimicrobial therapy, modulator treatment, and multidisciplinary follow-up have substantially improved survival, resulting in a growing population of adults living with CF. Although these advances may reduce hospitalization and intensive care utilization in some patients, adults with CF may still develop critical illness due to acute respiratory failure, severe infection, or complications of advanced multisystem disease. Therefore, as the adult CF population expands, understanding the characteristics and outcomes of those who require intensive care remains clinically relevant [6,7].

The management of critically ill patients with CF poses unique challenges. These patients frequently present with advanced lung disease, chronic colonization with multidrug-resistant pathogens, nutritional impairment, and extrapulmonary organ involvement, all of which may adversely affect ICU outcomes [8,9]. Previous studies from high-income countries have reported variable ICU mortality rates in CF, influenced by disease severity, age, and the presence of acute complications [10,11]. However, most available data originate from single-center cohorts or national registries in high-income settings, limiting the generalizability of findings to countries with different healthcare system structures [12]. Despite this extensive ICU infrastructure, nationwide data on the characteristics, distribution, and outcomes of critically ill patients with CF are scarce. In particular, little is known about where patients with CF receive intensive care, how ICU admissions are geographically distributed, and which clinical factors are associated with ICU mortality at the national level.

Therefore, the aim of this study was to provide the first nationwide descriptive analysis of ICU admissions among patients with cystic fibrosis in Türkiye over a six-year period. We sought to characterize the types and levels of ICUs to which patients with CF are admitted, to describe the geographic distribution of ICU admissions, and to compare clinical characteristics and outcomes between ICU survivors and non-survivors.

Materials and Methods

Study design and data source

This nationwide retrospective cohort study analyzed ICU admissions of adult patients with CF in Türkiye between January 1, 2018, and December 31, 2023. Data were retrieved from the Turkish Ministry of Health National Electronic Database, which prospectively collects administrative and clinical information from all public, university, and private hospitals across the country. Adult patients with cystic fibrosis were identified using ICD-10 code E84 and all of its subcategories recorded in the national database. Admission diagnoses and comorbid conditions were similarly identified using predefined ICD-10 codes. Admission diagnoses and comorbid conditions were similarly identified using predefined ICD-10 codes.

For each eligible patient, demographic characteristics, ICU admission data, primary admission diagnoses, comorbidities, ICU interventions, APACHE II scores, and ICU outcomes were extracted from the national database. The ICD-10 codes used for cystic fibrosis, admission diagnoses, and comorbid conditions are provided in Table 1.

Study population

Adult patients (≥18 years) with a diagnosis of cystic fibrosis who were admitted to an ICU between January 1, 2018, and December 31, 2023, were included. Cystic fibrosis was identified using ICD-10 code E84 and all corresponding subcategories. All ICU admissions meeting these criteria were included in the study. The unit of analysis was the individual patient, and the final cohort consisted of 126 unique adult patients with cystic fibrosis who were admitted to an ICU during the study period.

ICU characteristics and classification

ICUs were categorized as Level 1, Level 2, or Level 3 according to national regulatory definitions, which are based on staffing, monitoring capabilities, and the complexity of organ support that can be provided. ICU types were classified as mixed, anesthesiology, pulmonary, medical, surgical, and other specialized units. Information regarding the availability of each ICU level at individual hospitals was not available in the national dataset.

Table 1. ICD-10 codes used for identification of cystic fibrosis, admission diagnoses, and comorbidities.

Variable	ICD-10 code(s)
Cystic fibrosis	E84 and all subcategories
Respiratory failure	J96 and subcategories
Sepsis	A40–A41 and subcategories
Pulmonary embolism	I26 and subcategories
Acute kidney injury	N17 and subcategories
Liver failure	K72 and subcategories
Cerebrovascular disease	I60–I69
Gastrointestinal hemorrhage	K92.0–K92.2
Hypertension	I10–I15
Diabetes mellitus	E10–E14
Coronary artery disease	I20–I25
Heart failure	I50 and subcategories
Chronic kidney disease	N18 and subcategories

ICD-10, International Classification of Diseases, 10th Revision. Diagnostic categories include the listed parent codes and their corresponding subcategories unless otherwise specified.

Variables and outcomes

Demographic characteristics (age and sex), primary admission diagnoses, comorbid conditions, and ICU interventions—including non-invasive mechanical ventilation (NIMV), continuous renal replacement therapy (CRRT), and tracheotomy—were recorded. Disease severity was assessed using the Acute Physiology and Chronic Health Evaluation II (APACHE II) score at ICU admission. The primary outcome was ICU mortality. Secondary outcomes included ICU length of stay and the distribution of ICU admissions by region, ICU type, and care level.

Statistical analysis

Continuous variables were assessed for normality using the Kolmogorov–Smirnov test and are presented as medians with interquartile ranges (25th–75th percentiles). Categorical variables are summarized as frequencies and percentages. Comparisons between ICU survivors and non-survivors were performed using the Mann–Whitney U test for continuous variables and the chi-square test for categorical variables, as appropriate. All statistical tests were two-sided, and a p value <0.05 was considered statistically significant. Statistical analyses were conducted using IBM SPSS Statistics for Windows, version 23.0 (IBM Corp., Armonk, NY, USA).

This study was conducted in accordance with the principles of the Declaration of Helsinki. The analysis was based on anonymized, routinely collected national registry data, and individual patient consent was not required. The study protocol was approved by the Republic of Türkiye Ministry of Health with the approval number 95741342-020.

Results

ICU admissions and unit characteristics

Between 2018 and 2023, a total of 126 adult patients with cystic fibrosis were admitted to ICUs nationwide. Annual admissions ranged from 14 cases in 2018 and 2019 to 27 cases in 2020. No admissions occurred in Level 1 ICUs during the study period. Overall, 79 admissions (62.7%) were managed in Level 3 ICUs and 47 (37.3%) in Level 2 ICUs. Over time, admissions increasingly concentrated in Level 3 units. Regarding ICU type, mixed ICUs accounted for the largest proportion of admissions (38.8%), followed by anesthesiology ICUs (18.3%), pulmonary ICUs (13.5%), and medical ICUs (11.1%). Other specialized ICUs represented a small proportion of admissions (Table 2).

Table 2. Characteristics of intensive care units to which patients with CF were admitted.

Year n (%)	2018 14 (100)	2019 14 (100)	2020 27 (100)	2021 24 (100)	2022 23 (100)	2023 24 (100)	Total 126 (100)
Level of ICU							
Level 1	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Level 2	9 (64.3)	9 (64.3)	6 (22.2)	7 (29.2)	8 (34.8)	8 (33.3)	47 (37.3)
Level 3	5 (35.7)	5 (35.7)	21 (78.8)	17 (70.8)	15 (65.2)	16 (66.7)	79 (62.7)
ICU Type							
Mixed	8 (57.1)	5 (35.7)	15 (55.6)	6 (25.0)	7 (30.4)	8 (33.3)	49 (38.8)
Anesthesiology	1 (7.1)	3 (21.4)	6 (22.2)	6 (25.0)	6 (26.1)	1 (4.2)	23 (18.3)
Pulmonary	1 (7.1)	2 (14.3)	2 (7.4)	5 (20.8)	1 (4.4)	6 (25.0)	17 (13.5)
Medical	3 (21.4)	1 (7.1)	1 (3.7)	1 (4.1)	3 (13.0)	5 (20.8)	14 (11.1)
Surgical	1 (7.1)	1 (7.1)	1 (3.7)	1 (4.1)	0 (0)	4 (16.7)	8 (6.3)
Cardiovascular Surgery	0 (0)	1 (7.1)	0 (0)	2 (8.3)	1 (4.4)	0 (0)	4 (3.2)
Neurology	0 (0)	0 (0)	0 (0)	1 (4.1)	2 (8.7)	0 (0)	3 (2.4)
Thoracic Surgery	0 (0)	1 (7.1)	1 (3.7)	1 (4.1)	0 (0)	0 (0)	3 (2.4)
Coronary	0 (0)	0 (0)	0 (0)	1 (4.1)	1 (4.4)	0 (0)	2 (1.6)
Transplantation	0 (0)	0 (0)	0 (0)	0 (0)	1 (4.4)	0 (0)	1 (0.8)
Obstetric	0 (0)	0 (0)	0 (0)	0 (0)	1 (4.4)	0 (0)	1 (0.8)
Neurosurgical	0 (0)	0 (0)	1 (3.7)	0 (0)	0 (0)	0 (0)	1 (0.8)

ICU: intensive care unit, CF: cystic fibrosis

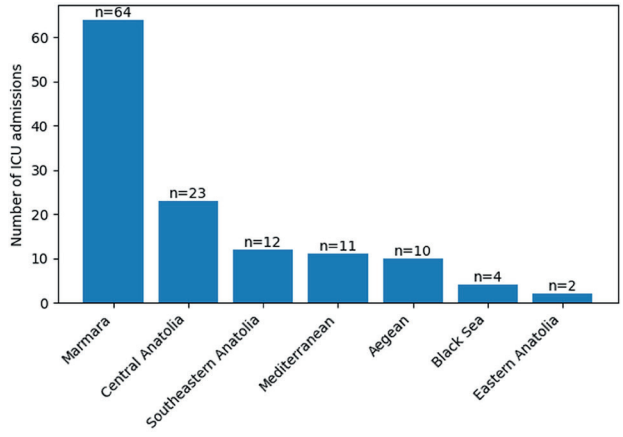


Figure 1. Geographic distribution of intensive care unit admissions among patients with cystic fibrosis (CF) across the seven regions of Türkiye.

Bars represent the number of intensive care unit (ICU) admissions in each region. Absolute numbers (n) are displayed above the bars to enhance transparency given the limited sample size.

Geographic distribution of ICU admissions

ICU admissions were unevenly distributed across the seven geographical regions of Türkiye (Figure 1). The majority of admissions occurred in the Marmara region (n=64, 50.8%), followed by the Central Anatolia region (n=23, 18.3%). Other regions contributed substantially fewer cases, including Southeastern Anatolia (n=12, 9.5%), Mediterranean (n=11, 8.7%), Aegean (n=10, 7.9%), Black Sea (n=4, 3.2%), and Eastern Anatolia (n=2, 1.6%).

Patient characteristics

The median age of the cohort was 30 years [23–41] and 60 patients (47.6%) were male. Overall ICU mortality was 32.5% (41 of 126 patients). When comparing survivors and non-survivors, non-survivors were older (median 32 [27–40] vs. 29 [21–46] years, p<0.01). Male sex was more frequent among survivors than non-survivors (54.1% vs. 34.1%, p=0.03). The most common

Table 3. Characteristics of critically Ill patients with cystic fibrosis.

	Total (n=126, 100%)	Survivors (n=85, 67.5%)	Non-survivors (n=41, 32.5%)	P value
Male gender, n (%)	60 (47.6)	46 (54.1)	14 (34.1)	0.03
Age *	30 [23-41]	29 [21-46]	32 [27-40]	<0.01
Admission Diagnosis, n (%)				
Respiratory failure	45 (35.7)	30 (35.3)	15 (36.6)	0.88
Sepsis	26 (20.6)	13 (15.3)	13 (31.7)	0.03
Pulmonary embolism	5 (4.0)	2 (2.4)	3 (7.3)	0.32
AKI	5 (4.0)	3 (3.5)	2 (4.9)	0.66
Liver failure	2 (1.6)	1 (1.2)	1 (2.4)	0.55
CVD	1 (0.8)	1 (1.2)	0 (0)	1.00
GIH	1 (0.8)	0 (0)	1 (2.4)	1.00
Comorbidities, n (%)				
Hypertension	46 (36.5)	22 (25.9)	24 (58.5)	<0.01
Diabetes mellitus	40 (31.8)	24 (28.2)	16 (39.0)	0.22
CAD	30 (23.8)	12 (14.1)	18 (43.9)	<0.01
CVD	24 (19.1)	11 (12.9)	13 (31.7)	0.01
Heart failure	11 (8.7)	4 (4.7)	7 (17.1)	0.04
CKD	10 (7.4)	5 (5.9)	5 (12.2)	0.29
APACHE II score*	12 [4-22]	7 [3-13]	21 [12-28]	<0.01
Non-invasive mechanical ventilation requirement , n (%)	74 (58.7)	53 (62.3)	21 (51.2)	0.23
Continous renal replacement therapy requirement, n (%)	10 (7.9)	8 (9.4)	2 (4.9)	0.49
Tracheotomy requirement, n (%)	12 (9.5)	7 (8.2)	5 (12.2)	0.52
Duration of ICU stay, days*	2 [1-7]	2 [1-7]	2 [1-5]	<0.01

* median [IQR],
AKI:acute kidney injury, CVD:cerebrovascular disease, GIH:gastrointestinal hemorrhage, CAD:coronary artery disease, CKD:chronic kidney disease, APACHE:acute physiology and chronic health evaluation, ICU:intensive care unit.

primary admission diagnosis was respiratory failure (35.7%), followed by sepsis (20.6%). The distribution of respiratory failure did not differ significantly between survivors and non-survivors (p=0.88). In contrast, sepsis was more frequent among non-survivors (31.7% vs. 15.3%, p=0.03). Hypertension (36.5%), diabetes mellitus (31.8%), and coronary artery disease (23.8%) were the most common comorbidities. Hypertension, coronary artery disease, cerebrovascular disease, and heart failure were significantly more frequent among non-survivors (all p<0.05) (Table 3).

The median APACHE II score at ICU admission was 12 [4–22]. Non-survivors had significantly higher APACHE II scores compared with survivors (21 [12–28] vs. 7 [3–13] , p<0.01). NIMV was required in 58.7% of patients and did not differ significantly between groups (p=0.23). CRRT and tracheotomy rates were low and were not significantly associated with ICU mortality. The median ICU length of stay was 2 days [1–7] in the overall cohort. Although the median duration was similar between survivors and non-survivors, the overall distribution of ICU length of stay differed significantly between the groups (p<0.01), reflecting differences in the distribution of values beyond the median alone (Table 3).

Discussion

In this nationwide retrospective cohort study, we describe the characteristics and outcomes of adult patients with CF admitted to ICUs across Türkiye over a six-year period. Several important findings emerged. First, ICU admissions were predominantly concentrated in Level 3 and mixed ICUs. Second, admissions demonstrated marked geographic clustering, with more than half occurring in the Marmara region. Third, overall ICU mortality was 32.5%. Mortality was significantly associated with older age, sepsis, cardiovascular comorbidities, and higher APACHE II scores at ICU admission. These findings provide the first national overview of ICU utilization and short-term outcomes among patients with CF in Türkiye. The predominance of Level 3 ICU admissions should be interpreted with caution. This distribution may reflect differences in local ICU infrastructure and referral pathways rather than illness severity alone. In addition, because cystic fibrosis is a relatively uncommon and specialized condition, clinicians may preferentially admit these patients to tertiary-level units for closer monitoring and access to advanced respiratory and multidisciplinary support, even when the initial degree of physiological instability is not severe.

ICU outcomes in patients with CF have evolved considerably over the past three decades. Early reports from the 1990s suggested that ICU admission—particularly when invasive mechanical ventilation was required—was associated with extremely poor prognosis, leading some to question the appropriateness of aggressive life-support measures in advanced CF. In one of the earliest cohort studies from the United States, ICU survival among intubated CF patients was limited and strongly influenced by transplant eligibility [13]. These findings contributed to a historically pessimistic view of ICU care in CF. In contrast, the overall ICU mortality in our contemporary nationwide cohort was 32.5%, suggesting that short-term outcomes are substantially more favorable than those described in historical series.

However, more contemporary studies have demonstrated substantial improvements in ICU survival, reflecting advances in ventilatory strategies, antimicrobial therapy, multidisciplinary CF care, and access to lung transplantation [14,15]. In a French multicenter study evaluating adults with CF admitted to ICUs, overall mortality was lower than previously reported, although invasive mechanical ventilation remained a strong determinant of death [16]. Similarly, subsequent studies from the United Kingdom and North America confirmed that ICU outcomes depend largely on the precipitating cause of respiratory failure, with better survival observed in patients intubated for potentially reversible conditions such as pneumothorax or hemoptysis compared with severe infective exacerbations [17,18]. Our findings are broadly consistent with this contemporary literature: respiratory failure was the most common reason for ICU admission (35.7%), while sepsis accounted for 20.6% of admissions and was more frequent among non-survivors (31.7% vs. 15.3%). Taken together, these findings place our national cohort within the evolving contemporary experience of critical care in adults with CF, in which ICU mortality remains substantial but is no longer uniformly poor.

Large administrative analyses from the United States further demonstrated that outcomes for adults with CF requiring invasive mechanical ventilation have improved over time, although mortality remains substantial, particularly among older patients and those with multiple comorbidities [10]. More recent data suggest that even advanced support modalities, including extracorporeal membrane oxygenation (ECMO), may be successfully applied in selected patients, particularly as a bridge to transplantation [19]. Against this evolving background, the ICU mortality rate of 32.5% observed in our nationwide cohort is consistent with contemporary registry-based reports describing short-term mortality between 25% and 40% among adults with cystic fibrosis admitted for acute respiratory failure or severe systemic complications [20,21]. In our cohort, APACHE II scores were substantially higher among non-survivors than survivors, indicating an association between greater physiological derangement at ICU admission and mortality. However, this finding should not be interpreted as demonstrating that acute illness severity is independent of the underlying severity of cystic fibrosis, since patients with more advanced disease may experience more pronounced physiological deterioration during acute exacerbations. Moreover, given the relatively young age of our cohort, the age component of APACHE II was unlikely to contribute substantially to the observed score differences. Several important cystic fibrosis–specific variables were unavailable in the national database, including baseline pulmonary function, CFTR modulator use, chronic microbiological colonization, nutritional status, and lung transplantation eligibility. These factors reflect important dimensions of underlying disease severity

and may influence both the likelihood of ICU admission and subsequent outcomes. Their absence limits our ability to distinguish the contribution of acute critical illness from that of chronic CF disease burden and may also contribute to residual confounding in the observed associations with ICU mortality.

Sepsis was more frequent among non-survivors in our cohort, suggesting that infectious complications may be associated with adverse short-term outcomes in critically ill adults with CF. This finding is consistent with previous reports identifying severe pulmonary and systemic infections as important contributors to clinical deterioration in CF [22,23]. However, given the observational design and the absence of adjustment for potential confounders in the primary analysis, these findings should be interpreted as associations rather than evidence of a causal or independent effect.

Our cohort was older than the population described in the Turkish National Cystic Fibrosis Registry [24]. This difference should be interpreted in the context of the populations studied. Our cohort included only adults with cystic fibrosis requiring ICU admission and therefore represents a selected subgroup rather than the broader CF population included in the national registry. Patients requiring critical care may be older and have a greater cumulative disease burden than patients represented in disease-specific registries. Differences in population coverage between the national ICU database and the Turkish CF registry may also have contributed to the observed age discrepancy.

The prevalence of hypertension and coronary artery disease in our cohort was substantially higher than that reported in cystic fibrosis registries and other contemporary CF populations. In a large multinational analysis including 5,649 adults from the UK CF Registry and 6,265 adults from the TriNetX database, hypertension was present in only 2.2% and 11.8% of patients, respectively [25]. Similarly, the 2024 UK Cystic Fibrosis Registry reported hypertension in 3.3% of adults with CF [26]. Although the higher prevalence observed in our study may partly reflect the selected nature of an older, critically ill adult cohort, differences in cohort composition alone may not fully account for the magnitude of these findings. Because comorbidities were identified using routinely recorded ICD-10 codes in an administrative database, variation in coding practices and diagnostic misclassification may also have contributed to the observed prevalence estimates. Accordingly, these comorbidity data should be interpreted with caution and should not be considered directly comparable with prevalence estimates derived from dedicated cystic fibrosis registries.

Cardiovascular comorbidities were more frequent among non-survivors in our cohort; however, their overall prevalence was considerably higher than that reported in contemporary cystic fibrosis populations. Cardiovascular disease has historically been considered uncommon in CF, although its clinical relevance may increase as the CF population ages and develops conventional cardiovascular risk factors [27]. Ticona et al. similarly emphasized that cardiovascular disease remains relatively uncommon in people with CF but may become increasingly important with advancing age [28]. Therefore, the high prevalence of hypertension and coronary artery disease observed in our relatively young ICU cohort should not be attributed solely to the aging of the CF population. As these comorbidities were identified using routinely recorded ICD-10 codes in an administrative database, differences in coding practices and diagnostic misclassification may have contributed to the observed estimates. These findings should therefore be interpreted cautiously.

The distribution of ICU admissions across levels should be interpreted within the context of the Turkish critical care system. The absence of Level 1 admissions and the predominance of Level 3 care may not reflect patient acuity alone. Differences in the availability of ICU levels across hospitals, regional referral pathways, and the tendency to manage patients with a rare and complex disease such as cystic fibrosis in tertiary centers may all have influenced the observed distribution. In particular, patients may have been admitted directly to higher-level units for closer monitoring and access to advanced respiratory and multidisciplinary support. Because information on the availability of individual ICU levels at each hospital was not available in the national database, the relative contribution of infrastructure, referral practice, and clinical severity cannot be determined from our data.

The marked geographic concentration of ICU admissions in the Marmara region may reflect regional differences in population density, availability of specialized CF services, and tertiary referral patterns rather than true geographic variation in the burden of critical illness. Because center-level referral pathways and CF expertise

were not available in the database, this finding should be interpreted cautiously.

The median ICU length of stay was 2 days. However, the administrative dataset did not provide sufficiently granular information on discharge destination, inter-ICU transfer, or treatment trajectory to determine the reasons for this relatively short duration. Therefore, ICU length of stay should be interpreted descriptively rather than as a direct indicator of disease severity or efficiency of care.

Finally, it is noteworthy that, in contrast to some historical studies emphasizing invasive mechanical ventilation as a near-uniformly fatal event in advanced CF [13], contemporary evidence—including our national data—supports a more nuanced perspective. Although detailed information regarding invasive mechanical ventilation, including its frequency, duration, and associated outcomes, was also not available in our database, ICU admission in CF should not be viewed as uniformly futile; rather, prognosis appears to depend on age, baseline comorbidity burden, acute severity, and the reversibility of the precipitating event [29,30]. This shift in understanding mirrors the broader transformation of CF from a predominantly pediatric fatal disease to a chronic adult condition with increasingly complex systemic manifestations.

The principal strength of this study lies in its nationwide design, encompassing all ICU levels and hospital types across Türkiye over a six-year period. By capturing data from a centralized national database, this study provides a comprehensive overview of ICU admissions among adults with cystic fibrosis in a real-world healthcare setting. To our knowledge, this is the first study to characterize ICU utilization and short-term outcomes of patients with CF at a national level in Türkiye. The national database includes administrative and clinical records from public, university, and private hospitals throughout Türkiye, providing broad coverage of ICU care across different hospital types and levels. This nationwide scope reduces the potential for single-center or regional selection bias and supports the representativeness of our findings for adults with CF receiving intensive care in Türkiye. However, because case identification relied on routinely recorded ICD-10 codes rather than a dedicated CF registry, some degree of under-ascertainment remains possible. Accordingly, these findings should be considered representative of the ICU-treated adult CF population captured within the national healthcare database rather than of the entire adult CF population in Türkiye.

Nevertheless, several limitations should be considered. First, the retrospective design and reliance on ICD-10–based administrative coding for case identification introduce the possibility of misclassification or under-ascertainment. Although standardized coding enhances nationwide comparability, variations in documentation practices may have resulted in incomplete identification of eligible patients. Second, granular clinical information—such as baseline pulmonary function, microbiological profiles, transplant eligibility, CFTR modulator use, and genetic subtype—was not available in the registry dataset. Third, outcomes were limited to the ICU period, and long-term survival, rehospitalization, and post-discharge functional status could not be evaluated. Information on the availability of Level 1, Level 2, and Level 3 ICUs at individual hospitals was not available. Therefore, the observed distribution of admissions across ICU levels may partly reflect local ICU infrastructure and referral patterns rather than patient acuity alone. Finally, given the limited number of outcome events and the sparse distribution of several candidate covariates, together with uncertainty regarding the validity of some ICD-10–derived comorbidity variables, multivariable modelling was not performed because of concern for overfitting and unstable adjusted estimates. Accordingly, the observed associations should be interpreted as exploratory and unadjusted.

Despite these limitations, this study provides an important contemporary national perspective on ICU care in adults with cystic fibrosis. The findings suggest that short-term ICU mortality is primarily associated with acute disease severity and infectious complications, while the observed regional clustering likely reflects healthcare system organization and referral pathways rather than true epidemiologic variation. These data contribute to the limited global literature on critical care utilization in CF and underscore the need for prospective, multicenter investigations incorporating detailed clinical variables and longitudinal follow-up to better inform risk stratification and policy planning.

Conclusion

In conclusion, this nationwide study provides the first comprehensive overview of ICU admissions among

adults with CF in Türkiye. ICU mortality in this cohort was comparable to contemporary reports from high-income countries and was primarily associated with acute disease severity, infectious complications, and comorbidity burden. The marked regional clustering of admissions likely reflects healthcare system organization and tertiary referral patterns rather than true epidemiologic variation. These findings underscore that ICU care in CF should not be viewed as uniformly futile but rather as a context-dependent clinical decision shaped by patient age, acute physiologic status, and reversibility of the precipitating event. Future prospective, multicenter studies incorporating detailed clinical data and long-term outcomes are warranted to refine risk stratification and optimize critical care strategies for this evolving adult CF population.

Author contributions

Conception and design: M.Y., E.G.; Data acquisition: M.Y., E.G., R.B.; Data analysis: M.Y., R.B., S.B., A.T., N.A.; Data interpretation: M.Y.; Drafting of the manuscript: M.Y., E.G., R.B.; Critical revision of the manuscript: S.B., A.T., N.A. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Republic of Türkiye Ministry of Health (Date: 27.11.2019, protocol no:95741342-020). The analysis was based on anonymized, routinely collected national registry data, and individual patient consent was not required.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that during the preparation of this study, the following AI-assisted technology was used: ChatGPT, OpenAI on manuscript preparation and revision period (01.05.26/12.09.26). Extent of Use: ChatGPT was used solely for language editing and to improve the clarity and readability of the manuscript. All AI-assisted text was critically reviewed and revised by the authors, who take full responsibility for the final content. The authors confirm that they have critically reviewed and edited any AI-generated content and take full responsibility for the integrity, accuracy, and originality of the publication. The authors certify that the original human contribution is maintained and that AI-assisted tools are not listed or cited as authors.

References

[1] Elborn JS. Cystic fibrosis. Lancet 2016;388(10059):2519-31. **[Crossref]**

[2] Stephenson AL, Sykes J, Stanojevic S, et al. Survival comparison of patients with cystic fibrosis in canada and the united states: a population-based cohort study. Ann Intern Med 2017;166(8):537-46. **[Crossref]**

[3] Keogh RH, Seaman SR, Barrett JK, Taylor-Robinson D, Szczesniak R. Dynamic prediction of survival in cystic fibrosis: a landmarking analysis using uk patient registry data. Epidemiology 2019;30(1):29-37. **[Crossref]**

[4] Cystic Fibrosis Foundation. Dispelling misconceptions about cystic fibrosis. Available at: https://www.cff.org/intro-cf/dispelling-misconceptions-about-cystic-fibrosis (Accessed on Aug 18, 2026).

[5] Ademhan Tural D, Şişmanlar Eyüboğlu T, Cinel G, et al. Current status of cystic fibrosis in Türkiye: data from the national registry. Thorac Res Pract 2025;26(5):238-47. **[Crossref]**

[6] Nick JA, Chacon CS, Brayshaw SJ, et al. Effects of gender and age at diagnosis on disease progression in long-term survivors of cystic fibrosis. Am J Respir Crit Care Med 2010;182(5):614-26. **[Crossref]**

[7] Hayes D, Kopp BT, Preston TJ, et al. Survival of patients with cystic fibrosis on ECMO: analysis of the extracorporeal life support organization registry. Int J Clin Exp Med 2014;7(5):1370-2.

[8] Flume PA, O'Sullivan BP, Robinson KA, et al. Cystic fibrosis pulmonary guidelines: chronic medications for maintenance of lung health. Am J Respir Crit Care Med 2007;176(10):957-69. **[Crossref]**

[9] Bell SC, Mall MA, Gutierrez H, et al. The future of cystic fibrosis care: a global perspective. Lancet Respir Med 2020;8(1):65-124. **[Crossref]**

[10] Siuba M, Attaway A, Zein J, et al. Mortality in adults with cystic fibrosis requiring mechanical ventilation. Cross-sectional analysis of nationwide events. Ann Am Thorac Soc 2019;16(8):1017-23. **[Crossref]**

[11] Saydain G, Islam A, Afessa B, Ryu JH, Scott JP, Peters SG. Outcome of patients with idiopathic pulmonary fibrosis admitted to the intensive care unit. Am J Respir Crit Care Med 2002;166(6):839-42. **[Crossref]**

[12] Kerem E, Reisman J, Corey M, Canny GJ, Levison H. Prediction of mortality in patients with cystic fibrosis. N Engl J Med 1992;326(18):1187-91. **[Crossref]**

[13] Sood N, Paradowski LJ, Yankaskas JR. Outcomes of intensive care unit care in adults with cystic fibrosis. Am J Respir Crit Care Med 2001;163(2):335-8. **[Crossref]**

[14] MacKenzie T, Gifford AH, Sabadosa KA, et al. Longevity of patients with cystic fibrosis in 2000 to 2010 and beyond: survival analysis of the Cystic Fibrosis Foundation patient registry. Ann Intern Med 2014;161(4):233-41. **[Crossref]**

[15] Smith MA, McGarry ME, Ly NP, Zinter MS. Outcomes of children with cystic fibrosis admitted to PICUs. Pediatr Crit Care Med 2020;21(10):e879-87. **[Crossref]**

[16] Texereau J, Jamal D, Choukroun G, et al. Determinants of mortality for adults with cystic fibrosis admitted in Intensive Care Unit: a multicenter study. Respir Res 2006;7(1):14. **[Crossref]**

[17] Jones A, Bilton D, Evans TW, Finney SJ. Predictors of outcome in patients with cystic fibrosis requiring endotracheal intubation. Respirology 2013;18(4):630-6. **[Crossref]**

[18] Hayes D, McCoy KS, Whitson BA, Mansour HM, Tobias JD. High-risk age window for mortality in children with cystic fibrosis after lung transplantation. Pediatr Transplant 2015;19(2):206-10. **[Crossref]**

[19] Gibilaro JM, Keating C, Benvenuto L, et al. Survival in cystic fibrosis after acute respiratory failure supported by extracorporeal membrane oxygenation and/or invasive mechanical ventilation. J Cyst Fibros 2022;21(4):669-74. **[Crossref]**

[20] Sanders DB, Bittner RCL, Rosenfeld M, Hoffman LR, Redding GJ, Goss CH. Failure to recover to baseline pulmonary function after cystic fibrosis pulmonary exacerbation. Am J Respir Crit Care Med 2010;182(5):627-32. **[Crossref]**

[21] Cystic Fibrosis Foundation. Patient registry. Available at: https://www.cff.org/medical-professionals/patient-registry (Accessed on Feb 11, 2026).

[22] Chun SW, Somers ME, Burgener EB. Highly effective cystic fibrosis transmembrane conductance (regulator) modulator therapy: shifting the curve for most while leaving some further behind. Curr Opin Pediatr 2024;36(3):290-5. **[Crossref]**

[23] Britto CJ, Taylor-Cousar JL. Cystic fibrosis in the era of highly effective CFTR modulators. Clin Chest Med 2022;43(4):xiii-xvi. **[Crossref]**

[24] Çocuk Solunum Yolu Hastalıkları ve Kistik Fibrozis Derneği. Ulusal kistik fibrozis kayıt sistemi 2024 yılı verileri. Available at: https://www.kistikfibrozisturkiye.org/wp-content/uploads/2025/12/UKKS-29Aralik2025-03.pdf (Accessed on Aug 18, 2026).

[25] Frost F, Nazareth D, Fauchier L, et al. Prevalence, risk factors and outcomes of cardiac disease in cystic fibrosis: a multinational retrospective cohort study. Eur Respir J 2023;62(4):2300174. **[Crossref]**

[26] Cystic Fibrosis Trust. Registry annual report. Available at: https://www.cysticfibrosis.org.uk/about-us/uk-cf-registry/reporting-and-resources (Accessed on Aug 18, 2026).

[27] Shah PH, Lee JH, Salvi DJ, Rabbani R, Gavini DR, Hamid P. Cardiovascular system ınvolvement in cystic fibrosis. Cureus 2021;13(7):e16723. **[Crossref]**

[28] Ticona JH, Lapinel N, Wang J. Future comorbidities in an aging cystic fibrosis population. Life (Basel) 2023;13(6):1305. **[Crossref]**

[29] Marques LS, Boschiero MN, Sansone NMS, Brienze LR, Marson FAL. Epidemiological profile of hospitalized patients with cystic fibrosis in Brazil due to severe acute respiratory infection during the COVID-19 pandemic and a systematic review of worldwide COVID-19 in those with cystic fibrosis. Healthcare (Basel) 2023;11(13):1936. **[Crossref]**

[30] King CS, Brown AW, Aryal S, Ahmad K, Donaldson S. Critical care of the adult patient with cystic fibrosis. Chest 2019;155(1):202-14. **[Crossref]**

Acta Medica 2026; 57(3): 283-292 DOI: 10.32552/actamedica.2026.1285

acta medica RESEARCH ARTICLE

Improvement in geriatric syndromes following clinical response to endoscopic third ventriculostomy in idiopathic normal pressure hydrocephalus

Yasemin Polat[1], Ömer Özpişkin[2], Merve Güner[3], Eyüp Bayatlı[2], Onur Özgüral[2], Gökmen Kahiloğulları[2], Mustafa Ağahan Ünlü[2], Mert Eşme[1], Cafer Balcı[1], Meltem Gülhan Halil[1], Mustafa Cankurtaran[1], Burcu Balam Doğu[1]

[1]Division of Geriatrics, Department of Internal Medicine, Faculty of Medicine, Hacettepe University, Ankara, Türkiye
[2]Department of Neurosurgery, Faculty of Medicine, Ankara University, Ankara, Türkiye
[3]Division of Geriatrics, Department of Internal Medicine, Faculty of Medicine, İstinye University, İstanbul, Türkiye

Abstract

Objective: Idiopathic normal pressure hydrocephalus (iNPH) is a potentially reversible cause of functional and cognitive decline in older adults. Although endoscopic third ventriculostomy (ETV) has emerged as an alternative treatment option in selected patients, its multidimensional impact on geriatric syndromes remains unclear. This study aimed to evaluate changes in comprehensive geriatric assessment (CGA) parameters following ETV in patients with iNPH.

Materials and Methods: This prospective observational study included 14 patients aged ≥55 years with iNPH who underwent ETV. Patients were categorized according to clinical improvement based on the iNPH Grading Scale. CGA domains—including Katz Activities of Daily Living (ADL), Lawton–Brody Instrumental (ADL), Clinical Frailty Scale (CFS), SARC-F, Mini Nutritional Assessment–Short Form (MNA-SF), Yesavage Geriatric Depression Scale (GDS), Older People's Quality of Life Questionnaire (OPQOL), and the Quick Mild Cognitive Impairment Screen-Turkish Version (qMCI-TR) were evaluated preoperatively and at 1 and 3 months after ETV.

Results: Eight patients (57.1%) demonstrated clinical improvement after ETV. At baseline, the clinical improvement group had higher Katz ADL, IADL, and qMCI-TR scores than the without improvement group. At 1st month, the clinical improvement group showed significant improvements in Katz ADL (Δ=1.0, p=0.025), CFS (Δ=−1.0, p=0.011), SARC-F (Δ=−2.0, p=0.011), MNA-SF (Δ=1.0, p=0.011), and OPQOL (p=0.017). At 3rd month, further significant improvements were observed in CFS (Δ=−1.0, p=0.011), OPQOL (Δ=6.0, p=0.017), SARC-F (Δ=−2.5, p=0.011), and MNA-SF (Δ=1.0, p=0.016). In contrast, patients without clinical improvement showed no consistent favorable changes across CGA domains.

Conclusion: In older adults with iNPH, clinical response to ETV was associated with meaningful improvements across multiple geriatric domains, particularly frailty, sarcopenia risk, nutritional status, functional status, and quality of life. These findings suggest that the benefits of ETV may extend beyond the classical symptom triad and support the integration of CGA into the perioperative evaluation of patients with iNPH.

Keywords: endoscopic third ventriculostomy, frailty, geriatric syndromes, idiopathic normal pressure hydrocephalus, older adults

Corresponding author: Yasemin Polat • **Email:** dryaseminplt@gmail.com

Received: March 5, 2026 **Accepted:** July 1, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

Introduction

Idiopathic normal pressure hydrocephalus (iNPH) is a clinical syndrome characterized by the classic triad of gait disturbance, cognitive dysfunction, and urinary incontinence, accompanied by ventriculomegaly on neuroimaging in the absence of elevated cerebrospinal fluid (CSF) pressure [1]. Originally described by Hakim and Adams in 1965, iNPH is a potentially reversible cause of cognitive and gait impairment in older adults [2]. The diagnosis is typically supported by radiological evidence of ventricular enlargement disproportionate to cortical atrophy, along with clinical improvement following CSF diversion procedures [3]. According to the American–European guidelines, the prevalence of iNPH is reported to be approximately 3.7% in the general population and 8.9% among individuals aged 80 years and older [4]. The prevalence of iNPH increases with age, is most commonly observed in adults over 60 years, and affects both sexes equally. Evidence supports the effectiveness of surgical management in iNPH, primarily through ventriculoperitoneal shunt (VPS) placement or endoscopic third ventriculostomy (ETV) as the main therapeutic options [5]. Without treatment, iNPH can lead to significant neurological decline and an elevated risk of mortality [6].

ETV has regained prominence as an effective and less invasive approach for the management of obstructive hydrocephalus, standing alongside traditional shunting procedures. ETV is a minimally invasive neurosurgical procedure that restores CSF circulation by creating an opening in the floor of the third ventricle, thereby bypassing impaired CSF pathways [5]. Compared with VPS, ETV avoids permanent shunt dependency and shunt-related complications and has emerged as an alternative treatment option in selected patients with iNPH. However, clinical response to ETV is variable, particularly in older adults, and reliable predictors of postoperative outcomes remain insufficiently defined [7].

Given the central role of functional status and frailty in older adults, neglecting geriatric syndromes may lead to an incomplete understanding of treatment outcomes after ETV [8]. Frailty encompasses functional decline across multiple physiological systems, including neurodegeneration, sarcopenia, and cognitive impairment [9]. However, some of the most pronounced structural and functional deterioration occurs within the musculoskeletal system, adversely affecting balance, mobility, disability, and ultimately the ability to live independently [10]. Geriatric syndromes such as malnutrition and sarcopenia, which are closely associated with reduced physiological reserve, may adversely affect postoperative recovery and functional improvement [11]. Exploring how these syndromes interact with neurosurgical interventions could provide valuable insights for patient selection, perioperative risk stratification, and individualized treatment planning [12] . Previous studies have demonstrated that ETV can be safely performed in selected older adults; however, these studies primarily focused on procedural success and did not comprehensively evaluate geriatric syndromes or multidimensional outcomes [13]. While the impact of geriatric syndromes on clinical outcomes has been explored in patients undergoing VPS, evidence specifically addressing how these syndromes influence outcomes after ETV remains limited. Therefore, the present study aimed to evaluate the impact of geriatric syndromes on clinical outcomes after ETV in patients with iNPH.

Material and Method

Study design and participants

This prospective observational study included 14 older adults with iNPH who underwent ETV at Ankara University Hospital between July 2023 and May 2024. Eligible patients were prospectively evaluated for clinical outcomes following surgery. Older adults aged ≥55 years with a diagnosis of iNPH, with symptom duration of at least 3 months, and CSF opening pressure between 70 and 200 mmH_2O were included in the study.

Based on postoperative clinical assessment, patients were categorized into two groups: those with clinical improvement and those without clinical improvement. Clinical improvement was determined using standardized clinical evaluations, including changes in gait, cognitive function, and urinary symptoms as assessed during follow-up visits according to the iNPH Grading Scale (iNPHGS) [14]. A ≥1-point improvement in the iNPHGS has been used in previous studies as an indicator of a favorable clinical response following surgical treatment and was therefore adopted as the criterion for clinical improvement in the present study [15,16].

Exclusion criteria were:

- Inability to cooperate
- Presence of active malignancy
- Secondary NPH
- Communicating hydrocephalus
- CSF pressure >200 mmHg
- Symptom duration <3 months
- Lack of informed consent

A total of 8 patients were excluded because of missing data or inability to complete follow-up assessments, as shown in Figure 1. Patients who met the inclusion criteria underwent a comprehensive geriatric assessment (CGA) as a baseline evaluation before ETV. Subsequent follow-up assessments, including repeat CGA, were conducted during routine postoperative visits at 1 and 3 months after the procedure. In the present study, the term "baseline" refers to the preoperative assessment performed prior to the ETV procedure.

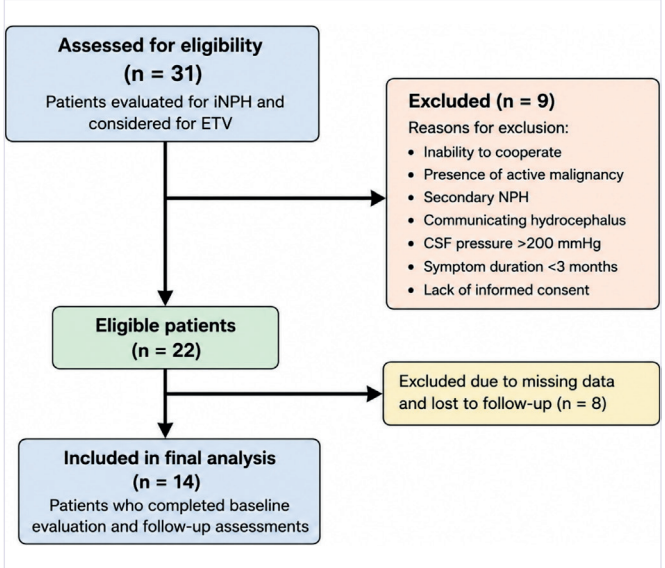


Figure 1. Flow diagram of patient selection, exclusions, and final study cohort.

Baseline demographic characteristics, clinical information on chronic diseases, and medication use were recorded. Laboratory parameters (electrolytes, thyroid-stimulating hormone, and complete blood count) were obtained from medical records. As part of the CGA, various screening and assessment tools were administered. These included the Katz Index of Activities of Daily Living (ADL), the Lawton–Brody Instrumental Activities of Daily Living (IADL) scale, the Quick Mild Cognitive Impairment Screen -Turkish Version (qMCI-TR), the Mini Nutritional Assessment–Short Form (MNA-SF), the Yesavage Geriatric Depression Scale (YGDS), the Clinical Frailty Scale (CFS), the Eating Assessment Tool-10 (EAT-10), SARC-F questionnaire, and the brief Older People's Quality of Life questionnaire (OPQOL-brief). All CGAs were conducted by the same experienced geriatrician throughout the study [17-31]. Quality of life was assessed using the OPQOL-brief, a validated short-form instrument specifically developed for older adults. The OPQOL-brief assesses multiple domains of quality of life, with higher scores indicating better perceived quality of life [32].

Statistical analysis

To evaluate the adequacy of the sample size, a sensitivity power analysis was performed using G*Power for a repeated-measures ANOVA (within–between interaction) with two groups and three measurements. With a total sample of 14 patients and α=0.05, the study had 80% power to detect a large effect size (f≈0.42). Therefore, the findings should be interpreted primarily in the context of large clinically meaningful effects. Data were analyzed using IBM SPSS Statistics for Windows, version 22.0 (IBM Corp., Chicago, IL, USA). The distributional properties of continuous variables were evaluated both visually (histograms and probability plots) and analytically using the Shapiro–Wilk and Kolmogorov–Smirnov tests. Continuous variables with normal distribution were presented as mean ± standard deviation (SD), whereas skew distributed variables were summarized as median [interquartile range(Q1-Q3) (IQR)]. Categorical data were expressed as a number (n) and a percentage (%). Between-group comparisons of categorical variables were performed using the Chi-square test or Fisher's exact test, as appropriate. For comparisons of continuous variables between the two independent groups, the independent samples t-test was used for normally distributed data, and the Mann–Whitney U test was used for non-normally distributed data. To evaluate within-group changes over time (baseline, 1st month, and 3rd month) and the group-by-time interaction effect between the two main groups, a repeated measures ANOVA was conducted for normally distributed variables. For non-normally distributed variables, within-group changes over time were assessed using the Friedman test with post-hoc Wilcoxon signed-rank tests A two-tailed p-value <0.05 was considered statistically significant.

Results

A total of 14 patients were evaluated, of whom 3 (21.4%) were female and 11 (78.6%) were male, with a mean age of 68.6 ± 8.4 years. Among the study cohort, 8 patients (57.1%) were classified as the clinical improvement group, while 6 (42.9%) formed the group without clinical improvement. Table 1 presents the clinical and demographic characteristics of the study population. The CGA findings of the study group at baseline and at the first and third months after ETV are shown in Table 2. No statistically significant changes were observed in CGA parameters, including Katz ADL, Lawton- Brody IADL, CFS, SARC-F, MNA-SF, GDS, OPQOL, and qMCI-TR across baseline, 1-month, and 3-month assessments in the overall study group (all p>0.05).

Table 1. The clinical and demographic features of the study population.

Features	n=14
Age, years	68.6 ±8.4
Sex, female	3 (21.4)
BMI, kg/m2	28.3 ±3.6
Education, 8 years and below	11 (78.6)
Marital Status, married	6 (42.9)
Smoking	4 (28.6)
Comorbidities	
Hypertension	7 (50.0)
Diabetes Mellitus	6 (42.9)
Coronary Artery Disease	7 (50.0)
Hyperlipidemia	4 (28.6)
Benign Prostate Hyperplasia	8 (57.1)
Clinical Improvement (+)	8 (57.1)

Variables are presented as n (%), mean ±SD, median (IQR) values. BMI: Body Mass Index.

Table 3 presents the baseline comprehensive geriatric and clinical characteristics of patients with iNPH stratified by clinical improvement status after ETV. At baseline, patients who achieved clinical improvement after ETV showed significantly higher Katz ADL, Lawton-Brody IADL, and qMCI-TR scores compared with those without improvement (p=0.015, p=0.024, and p=0.012, respectively). No significant differences were observed between the groups in age, sex, BMI, frailty status, nutritional status, or sarcopenia screening parameters (all p>0.05). Table 4 summarizes the changes in CGA scores at 1st month after ETV according to clinical improvement status. Statistically significant changes at 1st month after ETV were observed in the clinical improvement group for Katz ADL (median change: 1.0 [0.0 to 1.0], p = 0.025), CFS (median change: −1.0 [−1.0 to −1.0], p = 0.011), SARC-F (median change: −2.0 [−2.5 to −1.0], p = 0.011), MNA-SF (median change: 1.0 [1.0 to 1.0], p = 0.011), and OPQOL scores (median change: 3.0 [2.0 to 5.5], p = 0.017). No statistically significant changes were observed in the without clinical improvement group across CGA parameters (all p > 0.05). As shown in Table 5, at 3rd month after ETV, statistically significant

Table 2. Comprehensive geriatric assessment of the study group at baseline, first, and third months after treatment.

	Baseline (n=14)	1 month (n=14)	3 months (n=14)	p
Katz ADL	4.0 [3.0-5.0]	5.0 [2.0-6.0]	6.0 [1.0-6.0]	0.28
Lawton-Brody IADL	6.0 [2.0-7.0]	6.0 [2.0-7.0]	6.0 [2.0-8.0]	0.87
CFS	4.5 [4.0-5.0]	4.0 [2.0-6.0]	4.0 [2.0-7.0]	0.069
SARC-F	6.0 [4.0-6.0]	4.0 [2.0-8.0]	4.0 [2.0-8.0]	0.063
MNA-SF	9.0 [8.0-10.0]	10.0 [5.0-11.0]	10.0 [4.0-11.0]	0.61
Yesevage GDS	3.0 [2.0-4.0]	5.0 [1.0-12.0]	6.0 [1.0-12.0]	0.76
OPQOL	36 [26-42]	41.0 [20.0-46.0]	35.0 [13.0-52.0]	0.57
qMCI-TR	55.5 [38.0-59.0]	57.0 [17.5-59.25]	58.0 [18.0-64.0]	0.93

Variables are presented as n (%), mean ±SD, median (IQR) values. ADL: Activities of daily living; **CFS:** Clinical Frailty Scale; **IADL: Instrumental Activities of Daily Living;** MNA-SF: Mini-Nutritional Assessment Short-Form; GDS: Geriatric Depression Scale; OPQOL: Older People's Quality of Life questionnaire; qMCI-TR: the Quick Mild Cognitive Impairment Screen Turkish Version.

Table 3. Baseline demographic, laboratory values and clinical characteristics of iNPH patients with and without clinical improvement following ETV.

	Without clinical improvement (n=6)	With clinical improvement (n=8)	p
Age, years	72.0 [61.75-77.25]	68.5 [60.75-72.0]	0.44
Sex, female	-	3 (37.5)	0.21
BMI, kg/m2	25.7 [25.6-26.9]	28.9 [26.9-33.7]	0.053
Katz ADL	2.5 [1.0-4.0]	5.0 [4.0-6.0]	**0.015**
Lawton-Brody IADL	2.5 [2.0-6.0]	6.5 [6.0-8.0]	**0.024**
CFS	5.0 [5.0-6.0]	4.0 [3.5-4.5]	0.060
Living with Frailty (CFS≥4)	5 (83.3)	6 (75.0)	1.0
EAT-10	2.0 [0.0-6.0]	0.0 [0.0-0.0]	0.090
SARC-F	6.0 [6.0-7.0]	5.5 [3.5-6.0]	0.16
Sarcopenia Risk (SARC-F>4)	6 (100)	6 (75.0)	0.47
MNA-SF	8.0 [7.0-10.0]	9.0 [9.0-10.0]	0.39
Malnutrition (MNA-SF≤11)	5 (83.3)	8 (100.0)	0.43
Yesevage GDS	4.0 [3.0-10.0]	2.5 [2.0-3.5]	0.10
OPQOL	27.5 [25.0-37.0]	41.5 [35.0-45.0]	0.053
qMCI-TR	38.0 [29.0-46.0]	58.5 [55.25-62.25]	**0.012**
Degree of Improvement (iNPHGS)	0.0 [0.0-0.0]	2.0 [1.5-2.0]	**0.001**
Laboratory Values			
Na (mmol/L)	138 [136-139]	138 [136.5-140]	0.693
K (mmol/L)	4.4 [4.1-4.6]	4.1 [3.8-4.4]	0.364
AST (U/L)	17 [16-21]	21 [17.5-25.5]	0.300
ALT (U/L)	24 [9.0-26.0]	17.0 [13.5-23.5]	0.698
Creatinine (mg/dL)	1.05 [0.67-1.12]	0.78 [0.72-0.93]	0.519
TSH (mIU/L)	0.9 [0.9-1.7]	1.3 [1.1-3.5]	0.461
Ferritin (ng/mL)	7.7 [6.5-16.2]	7.3 [2.8-11.7]	0.439
Hb (g/dL)	12.1 [11.7-13.4]	13.8 [12.9-14.9]	0.196

Variables are presented as n (%), mean ±SD, median (IQR) values. Bold indicates $p < 0.05$. ADL: Activities of daily living; ALT: Alanine aminotransferase; AST: Aspartate aminotransferase; BMI: Body Mass Index; CFS: Clinical Frailty Scale; EAT-10: Eating Assessment Tool-10; Hb: Hemoglobin; IADL: Instrumental Activities of Daily Living; MNA-SF: Mini-Nutritional Assessment Short-Form; GDS: Geriatric Depression Scale; OPQOL: Older People's Quality of Life questionnaire; qMCI-TR: the Quick Mild Cognitive Impairment Screen Turkish Version; TSH: Thyroid-stimulating hormone.

changes were observed in the clinical improvement group for the CFS (median change: −1.0 [−1.5 to −1.0], p = 0.011), OPQOL scores (median change: 6.0 [2.5 to 8.5], p = 0.017), SARC-F (median change: −2.5 [−4.0 to −1.0], p = 0.011), and MNA-SF (median change: 1.0 [1.0 to 2.0], p = 0.016). In the without clinical improvement group, statistically significant changes were observed for MNA-SF (median change: −3.0 [−3.0 to −1.0], p = 0.039) and GDS scores (median change: 5.5 [2.0 to 11.0], p = 0.028). Other CGA parameters did not show statistically significant changes (p > 0.05).

Table 4. Change in scores (delta scores) of CGA parameters 1 month after the treatment by clinical improvement status.

	Without clinical improvement (n=6)	P	With clinical improvement (n=8)	p
Katz ADL	0.0 [-1.0; 1.0]	0.71	1.0 [0.0; 1.0]	**0.025**
Lawton-Brody IADL	-0.5 [-1.0; 0.0]	0.26	0.0 [0.0; 1.0]	0.083
CFS	0.0 [-1.0; 2.0]	0.83	-1.0 [-1.0; -1.0]	**0.011**
OPQOL	-4.5 [-9.0; 7.0]	0.30	3.0 [2.0; 5.5]	**0.017**
SARC-F	0.5 [-2.0; 4.0]	0.46	-2.0 [-2.5; -1.0]	**0.011**
MNA-SF	-1.5 [-3.0; 0.0]	0.068	1.0 [1.0; 1.0]	**0.011**
Yesevage GDS	2.5 [0.0; 11.0]	0.10	-1.5 [-2.0; 2.5]	0.83
qMCI-TR	-12.0 [-23.0;3.0]	0.17	2.5 [-3.5;4.0]	0.62

Higher scores indicate better status for Katz ADL, IADL, and MNA-SF; higher scores indicate worse status for CFS, Yesevage GDS, and SARC-F. Variables are presented as n (%), mean ±SD, median (IQR) values. Bold indicates $p < 0.05$. ADL: Activities of daily living; BMI: Body Mass Index; CFS: Clinical Frailty Scale; IADL: Instrumental Activities of Daily Living; MNA-SF: Mini-Nutritional Assessment Short-Form; GDS: Geriatric Depression Scale; OPQOL: Older People's Quality of Life questionnaire; qMCI-TR: the Quick Mild Cognitive Impairment Screen Turkish Version.

Table 5. Change in scores (delta scores) of CGA parameters 3 months after the treatment by clinical improvement status.

	Without Clinical improvement (n=6)	p	With Clinical improvement (n=8)	p
Katz ADL	-1.5 [-2.0;0.0]	0.10	0.5 [0.0;1.5]	0.063
Lawton-Brody IADL	-1.0 [-2.0; -1.0]	0.084	0.0 [0.0;1.0]	0.10
CFS	2.0 [-1.0;3.0]	0.83	-1.0 [-1.5;-1.0]	**0.011**
OPQOL	-12.0 [-13.0;-2.0]	0.29	6.0 [2.5;8.5]	**0.017**
SARC-F	2.0 [0.0;4.0]	0.46	-2.5 [-4.0;-1.0]	**0.011**
MNA-SF	-3.0 [-3.0;-1.0]	**0.039**	1.0 [1.0;2.0]	**0.016**
Yesevage GDS	5.5 [2.0;11.0]	**0.028**	-1.5 [-2.0;0.0]	0.26
qMCI-TR	-20.0 [-24.25;-3.25]	**0.027**	3.0 [1.5;4.0]	**0.029**

Higher scores indicate better status for Katz ADL, IADL, and MNA-SF; higher scores indicate worse status for CFS, Yesevage GDS and SARC-F. Variables are presented as n (%), mean ±SD, median (IQR) values. Bold indicates $p < 0.05$. ADL: Activities of daily living; BMI: Body Mass Index; CFS: Clinical Frailty Scale; IADL: Instrumental Activities of Daily Living; MNA-SF: Mini-Nutritional Assessment Short-Form; GDS: Geriatric Depression Scale; OPQOL: Older People's Quality of Life questionnaire; qMCI-TR: the Quick Mild Cognitive Impairment Screen Turkish Version

Discussion

The present study aimed to explore the association between baseline geriatric characteristics and clinical outcomes following ETV and to evaluate changes in CGA parameters after the procedure in patients with iNPH.

One of the main findings of this study was that baseline CGA profiles were associated with clinical response after ETV. Patients who achieved clinical improvement had higher baseline functional and cognitive scores. In addition, patients who showed clinical improvement following ETV exhibited favorable changes across multiple geriatric domains, including functional status, frailty, sarcopenia risk, nutritional status, and quality of life. These findings suggest that ETV may be associated with improvements in several geriatric domains beyond the traditional clinical triad.

Frailty is a dynamic and potentially reversible geriatric syndrome characterized by reduced physiological reserve and increased vulnerability to stressors [33]. Previous studies have shown that frailty status may improve when underlying contributors, such as impaired mobility or chronic disease burden, are effectively addressed. In this context, our findings suggest that clinical improvement after ETV in patients with iNPH may be accompanied by favorable changes in frailty status, reflecting the modifiable nature of frailty in

response to improvement in the underlying neurological condition. However, given the observational nature and small sample size of the study, these findings should be interpreted with caution. Patients with clinical improvement after ETV showed statistically significant improvement in frailty status during follow-up. Although frailty in patients with iNPH/NPH has been predominantly studied in the context of VPS, our findings help address an important gap in the literature by examining changes in frailty status following ETV [34].

Previous studies examining the relationship between frailty and outcomes after VPS have reported heterogeneous results. However, the overall evidence suggests that frail patients tend to experience higher postoperative complication rates and derive less clinical benefit following VPS compared with non-frail individuals [34,35]. The observed improvement in frailty among patients who achieved clinical improvement after ETV suggests that frailty status may change alongside clinical response. These findings support the consideration of frailty assessment as part of the evaluation of outcomes following ETV in older adults. Therefore, these findings suggest that ETV may be considered in selected frail patients with iNPH, although larger studies are required to confirm its effectiveness and safety in this population.

Functional status is a core component of CGA and closely associated with frailty and overall clinical outcomes in older adults. In patients with iNPH, gait disturbance and reduced mobility are major contributors to functional dependence. Previous studies evaluating ETV outcomes have primarily focused on gait function, a key element of the clinical triad; notably, Sankey et al. reported meaningful improvements in functional gait outcomes following primary ETV in selected iNPH patients [36]. Building on these observations, the present study extends outcome assessment beyond gait by incorporating CGA-based domains. Patients who achieved clinical improvement after ETV also demonstrated improvements in functional status during follow-up. This observation may reflect a relationship between clinical response and functional outcomes in patients with iNPH.

Malnutrition and sarcopenia are highly prevalent and interrelated geriatric syndromes that contribute to functional decline, vulnerability to stressors, and adverse clinical outcomes in older adults. Studies specifically evaluating changes in nutritional status or sarcopenia following ETV in patients with iNPH are scarce. Malnutrition is defined as a state of deficient nutrient intake or assimilation leading to altered body composition and diminished physical and cognitive function, whereas sarcopenia is characterized by the progressive loss of skeletal muscle strength and performance [11,37]. In the present study, patients who achieved clinical improvement following ETV also demonstrated improvements in MNA-SF and SARC-F scores. These findings suggest a possible association between clinical response after ETV and changes in nutritional status and sarcopenia risk. Although the underlying mechanisms remain uncertain, improvements in mobility, physical performance, and functional independence after neurological recovery may contribute to these observations, as these factors are recognized determinants of nutritional intake and muscle preservation in older adults. Given the observational design and small sample size, these findings should be interpreted with caution. Nevertheless, they suggest that nutritional status and sarcopenia-related parameters may change alongside clinical improvement in selected patients with iNPH.

In patients with iNPH, ETV success is primarily evaluated based on improvements in cognition, gait and balance, and urinary continence, and most outcome measures used to define procedural success focus on these three domains. Previous evidence indicates that although advanced age is not an absolute contraindication for ETV, outcomes in older patients are more variable, highlighting the need for CGA to optimize patient selection [38]. In our study, patients with better baseline functional and cognitive performance tended to achieve better clinical outcomes after ETV. These findings may indicate that earlier identification and intervention, prior to significant functional and cognitive decline, could be associated with improved treatment response in patients with iNPH.

Quality of life is a key patient-centered outcome in older adults with iNPH, reflecting the multidimensional impact of gait disturbance, cognitive impairment, and functional dependency. The significant increase in OPQOL scores observed in the clinically improved group following ETV suggests that successful CSF diversion may translate into meaningful gains in perceived well-being. This improvement is likely mediated by enhanced mobility, greater functional independence, and partial cognitive recovery after neurological improvement.

Previous studies have similarly demonstrated that effective treatment of iNPH can lead to measurable improvements in health-related quality of life [39,40]. Taken together, these findings suggest that patient-reported outcome measures such as OPQOL may provide additional information regarding treatment outcomes in patients with iNPH.

This study has several limitations that should be acknowledged. First, the relatively small sample size limits the statistical power of the analyses and may restrict the generalizability of the findings. Larger, multicenter cohorts are needed to validate our observations. Second, the follow-up period was relatively short; therefore, causal inferences regarding the longitudinal relationship between ETV and changes in geriatric parameters cannot be firmly established. Studies with longer follow-up durations would help clarify the temporal dynamics and sustainability of the observed improvements. Third, our analysis focused exclusively on patients undergoing ETV. Inclusion of patients treated with VPS could have enabled a more comprehensive comparison of the relative advantages and disadvantages of the two surgical approaches. Another limitation of the present study is that the classification of clinical improvement was based on predefined clinical assessments; however, the lack of universally standardized criteria for defining ETV success may have introduced potential classification bias. In addition, a priori power analysis was not conducted prior to data collection, which limits the interpretability of the present findings. The study may have been underpowered to detect moderate effect sizes that could still be clinically meaningful, and non-significant results should not be interpreted as evidence of the absence of an effect. Furthermore, potential confounding factors, including age, sex, and comorbidity burden, were not evaluated in adjusted analyses. Given the small sample size, multivariable analyses were not feasible. Therefore, the independent contribution of these factors to the observed changes in geriatric outcomes could not be determined, and residual confounding cannot be excluded. Future studies with larger populations should address these factors in greater detail. A major strength of this study is its evaluation of geriatric syndromes in older adults undergoing ETV, helping to address a gap in the literature. Importantly, the prospective follow-up design enabled longitudinal evaluation of geriatric domains. Although ETV is an increasingly utilized surgical approach in selected patients with iNPH, evidence focusing on its broader geriatric implications remains limited. By evaluating frailty, nutritional status, sarcopenia screening, and quality of life in this context, our findings help shed light on the multidimensional clinical impact of ETV beyond traditional neurological outcomes and contribute to filling an important gap in the current literature.

In conclusion, in older adults with iNPH, clinical improvement following ETV was accompanied by meaningful gains across multiple geriatric domains, including frailty, functional status, nutritional status, sarcopenia screening, and quality of life. These findings suggest that the benefits of ETV may extend beyond the classical symptom triad and highlight the dynamic and potentially modifiable nature of geriatric syndromes in this population. Incorporating CGA into the perioperative evaluation of iNPH patients may enhance patient selection, prognostication, and postoperative monitoring.

Author contributions

Conception: Y.P., M.A.Ü., M.G.H., M.C., B.B.D.; Design: Y.P., B.B.D.; Data acquisition: Y.P., Ö.Ö., M.G., G.K.; Data analysis: Y.P., Ö.Ö., M.G., B.B.D.; Data interpretation: Y.P., Ö.Ö., M.G., O.Ö., B.B.D.; Drafting of the manuscript: E.B., M.A.Ü., M.E., M.G.H., B.B.D.; Critical revision of the manuscript: Y.P., E.B., O.Ö., G.K., C.B., M.C., B.B.D. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Ethical approval

This study was approved by the Hacettepe University Faculty of Medicine Non-Interventional Clinical Research Ethics Committee (Date: July 11, 2023, Decision/Protocol No: GO 23/562). Informed consent was obtained from all participants involved in this study.

Data availability statement

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

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that no generative AI or AI-assisted technologies were used in the writing or preparation of this study.

References

[1] Wang Z, Zhang Y, Hu F, Ding J, Wang X. Pathogenesis and pathophysiology of idiopathic normal pressure hydrocephalus. CNS Neurosci Ther 2020;26(12):1230-40. **[Crossref]**

[2] Hakim S, Adams RD. The special clinical problem of symptomatic hydrocephalus with normal cerebrospinal fluid pressure. Observations on cerebrospinal fluid hydrodynamics. J Neurol Sci 1965;2(4):307-27. **[Crossref]**

[3] Passos-Neto CEB, Lopes CCB, Teixeira MS, Studart Neto A, Spera RR. Normal pressure hydrocephalus: an update. Arq Neuropsiquiatr 2022;80(5 Suppl 1):42-52. **[Crossref]**

[4] Andersson J, Rosell M, Kockum K, Lilja-Lund O, Söderström L, Laurell K. Prevalence of idiopathic normal pressure hydrocephalus: a prospective, population-based study. PLoS One 2019;14(5):e0217705. **[Crossref]**

[5] Harbaugh TD, Stoltzfus MT, Hallan DR, Daggubati L, Rizk EB. Ventriculoperitoneal shunting versus endoscopic third ventriculostomy for the surgical management of ıdiopathic normal pressure hydrocephalus: a retrospective cohort analysis. Cureus 2025;17(2):e78347. **[Crossref]**

[6] Jaraj D, Wikkelsø C, Rabiei K, et al. Mortality and risk of dementia in normal-pressure hydrocephalus: a population study. Alzheimers Dement 2017;13(8):850-7. **[Crossref]**

[7] Greuter L, Schenker T, Guzman R, Soleman J. Endoscopic third ventriculostomy compared to ventriculoperitoneal shunt as treatment for idiopathic normal pressure hydrocephalus: a systematic review and meta-analysis. Br J Neurosurg 2024;38(6):1276-82. **[Crossref]**

[8] Roy JM, Kazim SF, Macciola D, et al. Frailty as a predictor of postoperative outcomes in neurosurgery: a systematic review. J Neurosurg Sci 2024;68(2):208-15. **[Crossref]**

[9] Pilotto A, Custodero C, Maggi S, Polidori MC, Veronese N, Ferrucci L. A multidimensional approach to frailty in older people. Ageing Res Rev 2020;60:101047. **[Crossref]**

[10] Polat Y, Yalcin A, Yazihan N, et al. The relationship between frailty and serum alpha klotho levels in geriatric patients. Arch Gerontol Geriatr 2020;91:104225. **[Crossref]**

[11] Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing 2019;48(1):16-31. **[Crossref]**

[12] Luciano MG, Williams MA, Hamilton MG, et al. A randomized trial of shunting for ıdiopathic normal-pressure hydrocephalus. N Engl J Med 2025;393(22):2198-209. **[Crossref]**

[13] Niknejad HR, Depreitere B, De Vleeschouwer S, Van Calenbergh F, van Loon J. Results of endoscopic third ventriculostomy in elderly patients ≥65 years of age. Clin Neurol Neurosurg 2015;130:48-54. **[Crossref]**

[14] Kubo Y, Kazui H, Yoshida T, et al. Validation of grading scale for evaluating symptoms of idiopathic normal-pressure hydrocephalus. Dement Geriatr Cogn Disord 2008;25(1):37-45. **[Crossref]**

[15] Chen J, He W, Zhang X, et al. Value of MRI-based semi-quantitative structural neuroimaging in predicting the prognosis of patients with idiopathic normal pressure hydrocephalus after shunt surgery. Eur Radiol 2022;32(11):7800-10. **[Crossref]**

[16] Shinoda N, Hirai O, Hori S, et al. Utility of MRI-based disproportionately enlarged subarachnoid space hydrocephalus scoring for predicting prognosis after surgery for idiopathic normal pressure hydrocephalus: clinical research. J Neurosurg 2017;127(6):1436-42. **[Crossref]**

[17] Lawton MP, Brody EM. Assessment of older people: self-maintaining and instrumental activities of daily living. Gerontologist 1969;9(3):179-86.

[18] Isik EI, Yilmaz S, Uysal I, Basar S. Adaptation of the lawton instrumental activities of daily living scale to Turkish: validity and reliability study. Ann Geriatr Med Res 2020;24(1):35-40. **[Crossref]**

[19] Rockwood K, Song X, MacKnight C, et al. A global clinical measure of fitness and frailty in elderly people. CMAJ 2005;173(5):489-95. **[Crossref]**

[20] Özsürekci C, Balcı C, Kızılarslanoğlu MC, et al. An important problem in an aging country: identifying the frailty via 9 point clinical frailty scale. Acta Clin Belg 2020;75(3):200-4. **[Crossref]**

[21] Guigoz Y, Lauque S, Vellas BJ. Identifying the elderly at risk for malnutrition. The mini nutritional assessment. Clin Geriatr Med 2002;18(4):737-757. **[Crossref]**

[22] Sarikaya D, Halil M, Kuyumcu ME, et al. Mini nutritional assessment test long and short form are valid screening tools in Turkish older adults. Arch Gerontol Geriatr 2015;61(1):56-60. **[Crossref]**

[23] Belafsky PC, Mouadeb DA, Rees CJ, et al. Validity and reliability of the Eating Assessment Tool (EAT-10). Ann Otol Rhinol Laryngol 2008;117(12):919-24. **[Crossref]**

[24] Malmstrom TK, Morley JE. SARC-F: a simple questionnaire to rapidly diagnose sarcopenia. J Am Med Dir Assoc 2013;14(8):531-2. **[Crossref]**

[25] Bahat G, Yilmaz O, Kılıç C, Oren MM, Karan MA. Performance of SARC-F in regard to sarcopenia definitions, muscle mass and functional measures. J Nutr Health Aging 2018;22(8):898-903. **[Crossref]**

[26] O'Connor MN, O'Sullivan D, Gallagher PF, Eustace J, Byrne S, O'Mahony D. Prevention of hospital-acquired adverse drug reactions in older people using screening tool of older persons' prescriptions and screening tool to alert to right treatment criteria: a cluster randomized controlled trial. J Am Geriatr Soc 2016;64(8):1558-66. **[Crossref]**

[27] Yavuz BB, Varan HD, O'Caoimh R, et al. Validation of the Turkish version of the quick mild cognitive impairment screen. Am J Alzheimers Dis Other Demen 2017;32(3):145-56. **[Crossref]**

[28] Ertan T, Bugay G, Eker E. Geriatrik depresyon ölçeği kısa versiyonlarının Türk yaşlı toplumunda geçerlik ve güvenilirliği: GDS-15, GDS-10, GDS-4. 5. Ulusal Geropsikiyatri Sempozyumu; 2004.

[29] 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. **[Crossref]**

[30] Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. Studies of illness in the aged. The index of ADL: a standardized measure of biological and psychosocial function. JAMA 1963;185:914-9. **[Crossref]**

[31] Arik G, Varan HD, Yavuz BB, et al. Validation of Katz index of independence in activities of daily living in Turkish older adults. Arch Gerontol Geriatr 2015;61(3):344-50. **[Crossref]**

[32] Bowling A, Hankins M, Windle G, Bilotta C, Grant R. A short measure of quality of life in older age: the performance of the brief Older People's Quality of Life questionnaire (OPQOL-brief). Arch Gerontol Geriatr 2013;56(1):181-7. **[Crossref]**

[33] Clegg A, Young J, Iliffe S, Rikkert MO, Rockwood K. Frailty in elderly people. Lancet 2013;381(9868):752-62. **[Crossref]**

[34] Koo AB, Elsamadicy AA, Renedo D, et al. Hospital frailty risk score predicts adverse events and readmission following a ventriculoperitoneal shunt surgery for normal pressure hydrocephalus. World Neurosurg 2023;170:e9-e20. **[Crossref]**

[35] Courville E, Rumalla K, Kazim SF, et al. Risk analysis index as a preoperative frailty tool for elective ventriculoperitoneal shunt surgery for idiopathic normal pressure hydrocephalus. J Neurosurg 2023;140(4):1110-6. **[Crossref]**

[36] Sankey EW, Jusué-Torres I, Elder BD, et al. Functional gait outcomes for idiopathic normal pressure hydrocephalus after primary endoscopic third ventriculostomy. J Clin Neurosci 2015;22(8):1303-8. **[Crossref]**

[37] Cederholm T, Jensen GL, Correia MITD, et al. GLIM criteria for the diagnosis of malnutrition - A consensus report from the global clinical nutrition community. J Cachexia Sarcopenia Muscle 2019;10(1):207-17. **[Crossref]**

[38] Mehta NH, Greenberg ABW, Kahle KT. The Efficacy of endoscopic third ventriculostomy for idiopathic normal pressure hydrocephalus. World Neurosurg 2023;179:158-66. **[Crossref]**

[39] Mori E, Ishikawa M, Kato T, et al. Guidelines for management of idiopathic normal pressure hydrocephalus: second edition. Neurol Med Chir (Tokyo) 2012;52(11):775-809. **[Crossref]**

[40] Kazui H, Miyajima M, Mori E, Ishikawa M, SINPHONI-2 Investigators. Lumboperitoneal shunt surgery for idiopathic normal pressure hydrocephalus (SINPHONI-2): an open-label randomised trial. Lancet Neurol 2015;14(6):585-94. **[Crossref]**

Acta Medica 2026; 57(3): 293-295 DOI: 10.32552/actamedica.2026.1314

acta medica

LETTER TO THE EDITOR

When can antiseizure medications be safely withdrawn after encephalitis?

Okan Sökmen[1]

[1]Department of Neurology, Atatürk Sanatorium Training and Research Hospital, Ankara, Türkiye

Dear Editor,

Post-encephalitic epilepsy (PEE) remains one of the most important long-term neurological complications after encephalitis and may significantly affect long-term functional outcome and quality of life. Although the reported incidence varies according to etiology and follow-up duration, the long-term risk of epilepsy after encephalitis is consistently higher than in the general population [1-3]. Earlier population-based studies showed that the long-term risk of epilepsy after viral encephalitis was markedly higher in patients with early seizures during the acute phase [2]. More recent studies have linked later epilepsy risk to acute symptomatic seizures (ASyS), epileptiform electroencephalography (EEG) abnormalities, structural brain injury, and autoimmune etiologies, although the strength of these associations varies across cohorts [3,4].

Although several risk factors for PEE have been identified, their implications for antiseizure medication (ASM) management remain less clear. The central clinical question is no longer whether encephalitis increases epilepsy risk, but how clinicians should decide when to start, continue, and discontinue ASM treatment. In practice, ASM initiation after ASyS is common, particularly in patients with cortical injury, recurrent seizures, or epileptiform EEG findings. Status epilepticus and recurrent ASyS are generally accepted indications for treatment because both are associated with greater seizure burden and higher risk of subsequent epilepsy [3,4]. Uncertainty persists, however, in patients with isolated or brief ASyS who show apparent clinical recovery.

This uncertainty partly reflects evolving concepts of epileptogenesis. ASyS were traditionally regarded as transient events occurring during acute brain injury. Current models instead suggest that epilepsy develops through a process involving neuroinflammation, neuronal injury, synaptic reorganization, and progressive network hyperexcitability [5]. From this perspective, ASyS may represent an early manifestation of epileptogenesis rather than purely transient acute seizures. Still, the presence of ASyS does not necessarily mean that all patients require prolonged ASM treatment. The main difficulty is identifying patients with ongoing seizure risk while avoiding unnecessary long-term treatment in lower-risk individuals.

Although no standardized duration has been established for ASM treatment after encephalitis, several recent studies support shorter treatment durations in selected lower-risk patients. In the prospective PROSE study, median ASM duration was approximately 3 months, and longer treatment was not associated with lower relapse risk despite frequent prolonged ASM use in clinical practice [6]. Similarly, Dhawan et al. reported comparable seizure recurrence rates between 4-week and 12-week ASM regimens in children with acute encephalitis syndrome [7]. Routine prolonged ASM treatment after encephalitis may therefore be unnecessary in many patients, although evidence specifically addressing ASM withdrawal after

Corresponding author: Okan Sökmen • **Email:** drokansokmen@gmail.com

Received: May 10, 2026 **Accepted:** August 11, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

encephalitis remains limited. Prolonged ASM exposure may also contribute to unnecessary adverse effects and treatment burden in lower-risk patients.

Still, these observations should not be generalized to all patients. Persistent epileptiform EEG abnormalities, temporal lobe injury, autoimmune encephalitis, recurrent ASyS, and status epilepticus may argue against early ASM withdrawal and support closer follow-up in selected higher-risk individuals. Rather than relying on arbitrary treatment durations alone, ASM withdrawal after encephalitis may depend more on follow-up EEG findings, structural imaging, and overall clinical course (Table 1).

Acute epileptiform abnormalities during encephalitis do not necessarily predict persistent seizure risk, and follow-up EEG may therefore be more informative for ASM withdrawal decisions. Sivaraju et al. demonstrated that persistent epileptiform abnormalities on repeat EEG were strongly associated with later epilepsy development, whereas resolution of EEG abnormalities was associated with lower long-term seizure risk [8]. Routine acute-phase EEG recordings may not fully reflect later seizure risk or evolving cortical irritability.

Neuroimaging findings may further refine this decision-making process. Structural brain injury, particularly when involving cortical or temporal regions, is generally considered a higher-risk feature after encephalitis, although reported predictors vary across studies [3,6]. Ongoing cognitive symptoms, fluctuating behavioral changes, or episodic neurological complaints may require closer EEG follow-up before ASM discontinuation.

Autoimmune encephalitis deserves particular consideration because seizure evolution may differ substantially from that observed in infectious encephalitis. In a recent long-term study, Jia et al. reported that nearly one-third of autoimmune encephalitis patients with acute seizures developed treatment-resistant PEE during follow-up [9]. Epileptiform discharges during the subacute phase were among the strongest predictors of later epilepsy development. Not all autoimmune encephalitis subtypes appear to behave similarly. Patients with leucine-rich glioma-inactivated 1 (LGI1) or N-methyl-D-aspartate receptor (NMDAR) encephalitis may become seizure-free after immunotherapy and may not always require chronic ASM treatment, whereas delayed immunotherapy, ongoing inflammatory activity or structural temporal injury may support prolonged treatment and closer follow-up in selected patients [4,9].

A uniform ASM strategy after encephalitis may be difficult to justify. Patients with isolated brief ASyS, normal follow-up EEG findings, absence of cortical lesions, and rapid neurological recovery may be candidates for earlier ASM withdrawal after reassessment, often within the first few months. In contrast, recurrent seizures, status epilepticus, persistent epileptiform EEG abnormalities, autoimmune encephalitis, and temporal or cortical injury may argue against early ASM withdrawal. In practice, ASM withdrawal decisions may be better guided by follow-up EEG findings, structural imaging, and clinical course rather than by predetermined treatment durations alone.

Table 1. A practical approach to ASM withdrawal decisions after encephalitis.

Clinical domain	Lower-risk features	Higher-risk features	Possible implication for ASM withdrawal
Acute symptomatic seizures (ASyS)	Single brief seizure	Recurrent seizures or status epilepticus	Recurrent seizures may argue against early withdrawal
Follow-up EEG	Resolution of epileptiform abnormalities	Persistent epileptiform abnormalities	Persistent abnormalities may support prolonged treatment or closer follow-up
Neuroimaging	No cortical or temporal injury	Cortical, limbic, or mesiotemporal lesions	Structural injury may favor delayed withdrawal
Etiology	Self-limited infectious encephalitis	Autoimmune encephalitis or ongoing inflammatory activity	Withdrawal decisions may require individualized reassessment
Clinical course	Stable neurological recovery	Persistent cognitive/ behavioral symptoms	Repeat EEG evaluation may be helpful before withdrawal

Abbreviations: ASM: antiseizure medication; ASyS: acute symptomatic seizures; EEG, electroencephalography.

After encephalitis, the key clinical question is often not whether ASM should be started, but when it can be safely discontinued. Follow-up EEG, structural imaging, seizure characteristics, and etiology may all help guide these decisions. A more individualized approach may help reduce unnecessary long-term ASM use while supporting closer monitoring in higher-risk patients. Prospective studies specifically addressing ASM withdrawal after encephalitis are still needed.

Author contributions

Conception and design: O.S.; Data acquisition: O.S.; Data analysis: O.S.; Data interpretation: O.S.; Drafting of the manuscript: O.S.; Critical revision of the manuscript: O.S. The author reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Conflict of interest

The author declares that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The author declares that this study received no funding.

Generative AI statement

The author declares that during the preparation of this study, the following AI-assisted technology was used: ChatGPT (OpenAI) on May 2026. Extent of Use: ChatGPT (OpenAI) was used in May 2026 solely for language editing and improving the clarity of the manuscript. The scientific content, interpretation, and conclusions were developed by the author. The author confirms that he/she has critically reviewed and edited any AI-generated content and takes full responsibility for the integrity, accuracy, and originality of the publication. The author certifies that the original human contribution is maintained and that AI-assisted tools are not listed or cited as authors.

References

[1] Wagner JN, Moser T, Gruber J, et al. Characteristics of and risk factors for epilepsy after autoimmune and infectious encephalitis. BMC Neurol 2026;26(1):133. **[Crossref]**

[2] Annegers JF, Hauser WA, Beghi E, Nicolosi A, Kurland LT. The risk of unprovoked seizures after encephalitis and meningitis. Neurology 1988;38(9):1407-10. **[Crossref]**

[3] Zhong R, Zhang X, Chen Q, Li M, Guo X, Lin W. Acute symptomatic seizures and risk of epilepsy in autoimmune encephalitis: a retrospective cohort study. Front Immunol 2022;13:813174. **[Crossref]**

[4] Steriade C, Britton J, Dale RC, et al. Acute symptomatic seizures secondary to autoimmune encephalitis and autoimmune-associated epilepsy: conceptual definitions. Epilepsia 2020;61(7):1341-51. **[Crossref]**

[5] Pitkänen A, Engel J. Past and present definitions of epileptogenesis and its biomarkers. Neurotherapeutics 2014;11(2):231-41. **[Crossref]**

[6] Herzig-Nichtweiß J, Salih F, Berning S, et al. Prognosis and management of acute symptomatic seizures: a prospective, multicenter, observational study. Ann Intensive Care 2023;13(1):85. **[Crossref]**

[7] Dhawan SR, Sahu JK, Singhi P, Sankhyan N, Jayashree M. Comparison of 4 weeks versus 12 weeks antiseizure medication for acute symptomatic seizures in children with acute encephalitis syndrome: an open-label, randomized controlled trial. Seizure 2021;92:182-8. **[Crossref]**

[8] Sivaraju A, Tao A, Jadav R, et al. Antiseizure medication withdrawal, risk of epilepsy, and longterm eeg trends in acute symptomatic seizures or epileptic eeg patterns. Neurol Clin Pract 2024;14(6):e200342. **[Crossref]**

[9] Jia L, Kim CY, Pleshkevich M, et al. Long-term seizure outcomes in autoimmune encephalitis. Neurohospitalist 2025;15(4):343-54. **[Crossref]**

Acta Medica 2026; 57(3): 296-297 DOI: 10.32552/actamedica.2026.1327

acta medica

LETTER TO THE EDITOR

Letter to editor: Presentation and clinical outcome of adolescents with graves' disease

İrem Şenoymak[1], Mustafa Can Şenoymak[2]

[1]Department of Family Medicine, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, University of Health Sciences, İstanbul, Türkiye
[2]Department of Endocrinology and Metabolism, Hitit University, Çorum Erol Olçok Training and Research Hospital, Çorum, Türkiye

Dear Editor,

We read with interest the recent study evaluating the clinical features, treatment approaches and remission rates in adolescents diagnosed with Graves' disease [GD] [1]. We congratulate the authors for sharing their experience with this rare condition, the clinical course of which can vary from patient to patient. The data on long-term antithyroid drug (ATD) treatment, remission rates and the need for definitive treatment provide valuable insights into the management of paediatric GD.

As the average follow-up period was only 13 months and remission could only be assessed in the eight patients who successfully completed ATD treatment, the reported remission rate of 15.7 per cent should be interpreted with caution. The low remission rate may reflect the limited follow-up period rather than a more aggressive disease course. A meta-analysis comprising 20 studies involving paediatric and adolescent patients supports this view: remission rates were found to be 15.5 per cent in the second year of ATD treatment, 24.1 per cent in the fourth year, and 33.0 per cent after the sixth year or later. Differences in treatment duration and follow-up periods were identified as the main causes of heterogeneity between the studies [2]. Continued follow-up of the current cohort will help to clarify whether the observed rate truly reflects the severity of the disease or is largely a result of the short observation period.

The sample size is limited for making statistical inferences. However, the association between thyroid volume SDS and recurrence is consistent with the findings of previous studies. It has been consistently reported that a smaller goitre volume and lower TRAb levels at the time of diagnosis are associated with a more favourable prognosis, whilst the presence of a goitre is an independent predictor of recurrence [3,4]. Furthermore, a more detailed evaluation of serial TRAb measurements during the follow-up period could have contributed to a better understanding of the factors influencing remission and recurrence in this cohort.

Notably, one in four patients in whom a thyroid nodule was detected was diagnosed with papillary thyroid carcinoma. The risk of malignancy in thyroid nodules in the paediatric age group is approximately five times higher than in adults [5]. However, differentiated thyroid carcinoma developing against a background of Graves' disease is quite rare and may present a diagnostic challenge for clinicians. As the diffuse thyroid hyperplasia and increased vascularisation characteristic of Graves' disease may make it difficult to detect malignant nodules on ultrasound, it is of great importance to maintain careful and systematic ultrasound monitoring in these patients.

In conclusion, this valuable case series clearly demonstrates that Graves' disease in adolescents requires long-term follow-up, that remission rates

Corresponding author: İrem Şenoymak • **Email:** ireemakman94@gmail.com
Received: June 19, 2026 **Accepted:** July 7, 2026 **Published online:** September 30, 2026

Copyright © 2026 The Author(s). Published by Hacettepe University Faculty of Medicine. This is an open access article distributed under the Creative Commons Attribution License (CC BY), which permits unrestricted use, distribution, and reproduction in any medium or format, provided the original work is properly cited.

— while lower than those observed in adults — may improve with extended treatment duration, and that concomitant thyroid nodules warrant careful evaluation. We believe that larger, long-term, multicenter studies will contribute to a better understanding of the factors predictive of remission.

Author contributions

Conception and design: İ.Ş., M.C.Ş.; Data acquisition: İ.Ş.; Data analysis: M.C.Ş.; Data interpretation: İ.Ş.; Drafting of the manuscript: İ.Ş.; Critical revision of the manuscript: İ.Ş., M.C.Ş. All authors reviewed the results, approved the final version of the manuscript, and agreed to be accountable for all aspects of this study.

Conflict of interest

The authors declare that this study was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors declare that this study received no funding.

Generative AI statement

The authors declare that during the preparation of this study, the following AI-assisted technology was used: ChatGPT-5 on 19/07/2026. Extent of Use: The content and interpretations of this study were developed by us; artificial intelligence (ChatGPT) was utilized solely for language translation, as well as to ensure academic English standards and fluency of expression. The scientific framework, analysis, and conclusions of the text are entirely the product of my own work and academic assessment. The authors confirm that they have critically reviewed and edited any AI-generated content and take full responsibility for the integrity, accuracy, and originality of the publication. The authors certify that the original human contribution is maintained and that AI-assisted tools are not listed or cited as authors.

References

[1] Ünsal Y, Özaltın N, Buluş AD. Presentation and clinical outcome of adolescents with graves' disease. Acta Medica 2025;56(4):246-52. **[Crossref]**

[2] Li Y, Wang XM, Shi WY, Chen JJ, Song YN, Gong CX. Effect of antithyroid drugs treatment duration on the remission rates of graves' disease in children and adolescents: a single-arm meta-analysis and systematic review. Clin Endocrinol (Oxf) 2025;102(2):196-204. **[Crossref]**

[3] Gastaldi R, Poggi E, Mussa A, et al. Graves disease in children: thyroid-stimulating hormone receptor antibodies as remission markers. J Pediatr 2014;164(5):1189-94.e1. **[Crossref]**

[4] Rho JG, Kum CD, Seo YJ, Shim YS, Lee HS, Hwang JS. Long-term outcomes of Graves' disease in children and adolescents receiving antithyroid drugs. Ann Pediatr Endocrinol Metab 2021;26(4):266-71. **[Crossref]**

[5] Kanna E, Panayiotou D, Ganaiem A, et al. Papillary thyroid carcinoma in a pediatric patient with Graves' disease: a case report and literature review. Cureus 2026;18(3):e104829. **[Crossref]**