E-ISSN: 2547-9431

BAGCILAR MEDICAL BULLETIN

Bağcılar Tıp Bülteni

Volume 11, Issue 3, September 2026

ORIGINAL RESEARCHES / ARAŞTIRMALAR

Gabapentinoids and Vitamin Loss
Deniz Öke, Zeynel Karakullukçuoğlu

Vaccination and ICU Outcomes in COVID-19
Sezen Kumaş Solak, Ceren Gür, Burcu Akyüz, Mehmet Salih Sevdi, Serdar Demirgan, Rasim Onur Karaoğlu, Ayşin Selcan

Maternal IFN α-1 in Early-onset Preeclampsia
Barış Boza, Verda Alpay, Fırat Ersan, Hümeyra Öztürk Emre

AI Models in Scabies Information Delivery
Orhan Şen

Admission Respiratory Rate in Pregnancy ICU
Karolin Ohanoğlu Çetinel, Gülçin Hilal Alay, Emilya Zeynalli, Tasnim Sabouni, Rümeysa Akpınar, İsmail Alay

Inflammatory Markers in Obese Children
Aysun Boğa, Ezgi Nafile Sayman, Sümeyra Gedik Çalışkan, Feyza Nur Akan Çelen, Elif Betül Keskin Çetinkaya, Yeşim Coşkun, Şirin Güven

Regional Anesthesia on YouTube
Merve Bulun Yediyıldız, Kübra Taşkın, İrem Durmuş, Hülya Yılmaz Ak, Barış Sandal, Duygu Demiriz Gülmez

Shock Index Derivatives and Burn Mortality
Muhammed Koru, Emin Uysal, Burak Demirci, Abuzer Coşkun, Mücahit Yalçınkaya

Hemoptysis Recurrence After Bronchial Artery Embolization
Eda Cingöz, Ali Dablan, Mustafa Fatih Arslan, Oğuzhan Türksayar, Mehmet Cingöz, Çağrı Erdim, İlhan Nahit Mutlu, Tevfik Güzelbey, Özgür Kılıçkesmez

Anemia and Hypoalbuminemia in Acute HF
Can Baba Arın, Mehmet Zafer Aydın, Mehmet Kamil Teber, Nizamettin Selçuk Yelgeç

Paediatric Drowning Cases
Emre Aygün, Mustafa Safa Kasım, Mustafa Oğur, Ercüment Petmezci, Nazan Dalgıç

Tumor Localization and Breast Cancer Presentation
Erdoğan Şeyran, Azer Gökmen

Disability in Adults Aged ≥80 Years
Esra Çataltepe, Ayşe Fadıloğlu, Eda Çeker

Accidental Poisoning in the PICU
Kübra Boydağ Güvenç, Ebru Güney Şahin, İdris Abdullah Yılmaz, Fatih Varol, Cansu Durak

Sarcopenia Risk Perception in Parkinson's Disease
Hatice Betigül Meral, Abdülkadir Ermiş, Aygül Tantik Pak, Mehmet Yalman, Eren Toplutaş, Ece Akın, Sena Tolu

Surgical Outcomes of Gastric GISTs
Nurettin Şahin, Turgut Dönmez, Mahmut Said Değerli, Ferman Tevfik Özyalvaç, Süleyman Büyükaşık, Allanur Mametkulyyev, Duhan Kanmaz, Ahmet Sürek

CASE REPORT / OLGU SUNUMU

Occult Bronchial Carcinoid Tumor
Hülya Abalı, Ayşenur İlkyaz, Çiğdem Sabancı, Sida Gösterici, Fatma Tokgöz Akyıl

Periodical Publication of University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital

BAGCILAR MEDICAL BULLETIN
Bağcılar Tıp Bülteni

Publisher Contact
Address: Molla Gürani Mah. Kaçamak Sk. No: 21/1 34093 İstanbul, Turkey
Phone: +90 (530) 177 30 97
E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr **Publisher Certificate Number:** 14521

Publication Date: September 2026
E-ISSN: 2547-9431
International scientific journal published quarterly.

ADVISORY BOARD / DANIŞMA KURULU

ADVISORY BOARD / DANIŞMA KURULU

ADVISORY BOARD / DANIŞMA KURULU

DANIŞMA KURULU / ADVISORY BOARD

DANIŞMA KURULU / ADVISORY BOARD

DANIŞMA KURULU / ADVISORY BOARD

Please refer to the journal's webpage **(https://www.behmedicalbulletin.org/)** for "Journal Policies", and "Instructions to Authors" .

The editorial and publication process of the Bağcılar Medical Bulletin are shaped in accordance with the guidelines of ICMJE, WAME, CSE, COPE, EASE, and NISO. Bağcılar Medical Bulletin is indexed in **TÜBİTAK/ULAKBİM, EBSCO (Central & Eastern European Academic Source), Gale, Embase, Türk Medline, Turkey Citation Index, ProQuest, J-Gate, DOAJ** and **ScopeMed**.

The journal is published online.

Owner: In the name of T.C. Health Ministry, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Hospital Administrator Assoc. Prof., Aykut Çolakerol (University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Clinic of Urology, İstanbul, Turkey).

Responsible Manager: Assoc. Prof. Murat Altuntaş

Derginin "Dergi Politikaları"ve "Yazarlara Talimatlar" konularında bilgi almak için lütfen web sayfasına **(https://www.behmedicalbulletin.org/)** başvurun.

Bağcılar Tıp Bülteni ve/veya editörleri, ICMJE, COPE, WAME, CSE ve EASE üyeleridir ve bu kuruluşların önerilerini takip ederler. Bağcılar Tıp Bülteni, **TÜBİTAK/ULAKBİM, EBSCO (Central & Eastern European Academic Source), Gale, Embase, Türk Medline, Türkiye Atıf Dizini, ProQuest, J-Gate, DOAJ** ve **ScopeMed** dizinlerde indekslenmiştir.

Dergi, çevrimiçi olarak yayımlanmaktadır.

Sahip: T.C. Sağlık Bakanlığı, Sağlık Bilimleri Üniversitesi, İstanbul Bağcılar Eğitim ve Araştırma Hastanesi adına Başhekim Doç. Dr. Aykut Çolakerol (Sağlık Bilimleri Üniversitesi, İstanbul Bağcılar Eğitim ve Araştırma Hastanesi, Üroloji Kliniği, İstanbul, Türkiye).

Sorumlu Yönetici: Doç. Dr. Murat Altuntaş

BAGCILAR MEDICAL BULLETIN
Bağcılar Tıp Bülteni

CONTENTS / İÇİNDEKİLER

ORIGINAL RESEARCHES / ARAŞTIRMALAR

279 **Retrospective Analysis of Vitamin B12 and Folic Acid Levels Before and After Antiepileptic (Anticonvulsant) Treatment in Patients with Neuropathic Pain**
Nöropatik Ağrı Nedeniyle Antiepileptik (Antikonvülzan) Tedavi Alan Hastalarda B12 ve Folik Asit (Folat) Düzeyinin Tedavi Öncesi ve Sonrası Retrospektif İncelenmesi
Deniz Öke, Zeynel Karakullukçuoğlu; İstanbul, Turkey

290 **Impact of COVID-19 Vaccination Status on Intensive Care Unit Mortality in Patients with COVID-19 Pneumonia: A Retrospective Observational Study**
COVID-19 Pnömonisi Tanısıyla Yoğun Bakım Ünitesinde Yatan Hastalarda Aşılanma Durumunun Mortalite Üzerine Etkisi: Retrospektif Gözlemsel Çalışma
Sezen Kumaş Solak, Ceren Gür, Burcu Akyüz, Mehmet Salih Sevdi, Serdar Demirgan, Rasim Onur Karaoğlu, Ayşin Selcan; İstanbul, Turkey

297 **Maternal Immune Dysregulation in Preeclampsia: A Comparative Analysis of Maternal Serum Interferon Alpha-1 (IFN α-1) Concentrations in Preeclamptic and Healthy Pregnancies**
Preeklampside Maternal İmmün Disregülasyon: İnterferon Alfa-1 (IFN α-1) Maternal Serum Konsantrasyonlarının Preeklamptik ve Sağlıklı Gebeliklerde Karşılaştırmalı Analizi
Barış Boza, Verda Alpay, Fırat Ersan, Hümeyra Öztürk Emre; İstanbul, Turkey

307 **Comparative Evaluation of Artificial Intelligence Models in Providing Patient-Oriented Information on Scabies: Implications for Public Health and Clinical Practice**
Uyuz Hastalığı Hakkında Hasta Odaklı Bilgi Sağlamada Yapay Zeka Modellerinin Karşılaştırmalı Değerlendirmesi: Halk Sağlığı ve Klinik Uygulama Açısından Çıkarımlar
Orhan Şen; Kayseri, Turkey

319 **Respiratory Rate at Admission as an Independent Predictor of Intensive Care Requirement in Pregnant Women with COVID-19: A Retrospective Cohort Study**
Başvuru Sırasındaki Solunum Sayısının COVID-19'lu Gebe Kadınlarda Yoğun Bakım Gereksiniminin Bağımsız Bir Prediktörü Olarak Değerlendirilmesi: Retrospektif Kohort Çalışması
Karolin Ohanoğlu Çetinel, Gülçin Hilal Alay, Emilya Zeynalli, Tasnim Sabouni, Rümeysa Akpınar, İsmail Alay; İstanbul, Turkey

327 **Association Between Obesity Severity and Inflammatory Markers in Obese Children**
Obez Çocuklarda Obezite Şiddeti ile Enflamatuvar Belirteçler Arasındaki İlişki
Aysun Boğa, Ezgi Nafile Sayman, Sümeyra Gedik Çalışkan, Feyza Nur Akan Çelen, Elif Betül Keskin Çetinkaya, Yeşim Coşkun, Şirin Güven; İstanbul, Turkey

333 **Regional Anesthesia on YouTube: Are the Most Viewed Videos the Most Valuable? A Cross-sectional Content Analysis**
YouTube'da Rejyonel Anestezi: En Çok İzlenen Videolar En Değerli Olanlar mı? Kesitsel Bir İçerik Analizi
Merve Bulun Yediyıldız, Kübra Taşkın, İrem Durmuş, Hülya Yılmaz Ak, Barış Sandal, Duygu Demiriz Gülmez; İstanbul, Turkey

341 **Predictive Performance of Shock Index Derivatives for Mortality in Hospitalized Adult Burn Patients Presenting to the Emergency Department**
Acil Servise Başvuran ve Hastaneye Yatırılan Erişkin Yanık Hastalarında Mortalitenin Öngörülmesinde Şok İndeksi Türevlerinin Prediktif Performansı
Muhammed Koru, Emin Uysal, Burak Demirci, Abuzer Coşkun, Mücahit Yalçınkaya; Tokat, İstanbul, Manisa, Turkey

350 **Factors Associated with the Recurrence of Hemoptysis Following Bronchial Artery Embolization**
Bronşiyal Arter Embolizasyonu Sonrasında Nüks Hemoptizi ile İlişkili Faktörler
Eda Cingöz, Ali Dablan, Mustafa Fatih Arslan, Oğuzhan Türksayar, Mehmet Cingöz, Çağrı Erdim, İlhan Nahit Mutlu, Tevfik Güzelbey, Özgür Kılıçkesmez; İstanbul, Turkey

BAGCILAR MEDICAL BULLETIN
Bağcılar Tıp Bülteni

CONTENTS / İÇİNDEKİLER

359 **Prognostic Impact of Combined Anemia and Hypoalbuminemia Phenotypes in Patients Hospitalized with Acute Heart Failure**
Akut Kalp Yetersizliği Hastalarında Kombine Anemi ve Hipoalbüminemi Fenotiplerinin Prognostik Etkisi
Can Baba Arın, Mehmet Zafer Aydın, Mehmet Kamil Teber, Nizamettin Selçuk Yelgeç; İstanbul, Kayseri, Turkey

366 **Investigation of Childhood Drowning Cases Presenting to the Emergency Department of a Tertiary Care Paediatric Hospital Near the Sea Coast**
Deniz Kıyısına Yakın Bir Üçüncü Basamak Pediyatri Hastanesinin Acil Servisine Başvuran Çocuk Boğulma Olgularının İncelenmesi
Emre Aygün, Mustafa Safa Kasım, Mustafa Oğur, Ercüment Petmezci, Nazan Dalgıç; İstanbul, Turkey

376 **Association of Tumor Distribution Pattern with Stage at Diagnosis and Presentation Pattern in Patients with Invasive Breast Cancer**
İnvaziv Meme Kanserli Hastalarda Tümör Dağılım Paterni ile Tanı Anındaki Evre ve Başvuru Şekli Arasındaki İlişki
Erdoğan Şeyran, Azer Gökmen; Van, Turkey

387 **Factors Associated with Disability in Community-dwelling Oldest-old: Insights from Comprehensive Geriatric Assessment-based Nomogram**
Seksen Yaş ve Üzeri Toplumda Yaşayan Yaşlılarda Disabilite ile İlişkili Faktörler: Kapsamlı Geriatrik Değerlendirmeye Dayalı Nomogram
Esra Çataltepe, Ayşe Fadıloğlu, Eda Çeker; Konya, Osmaniye, Ankara, Turkey

396 **Clinical Profile and Advanced Treatment Requirements in Children Admitted to the Pediatric Intensive Care Unit for Accidental Poisoning**
Çocuk Yoğun Bakım Ünitesine Yatırılan Kazara Zehirlenme Olgularında Klinik Profil ve İleri Tedavi Gereksinimi
Kübra Boydağ Güvenç, Ebru Güney Şahin, İdris Abdullah Yılmaz, Fatih Varol, Cansu Durak; İstanbul, Turkey

404 **Sarcopenia Risk Perception in Parkinson's Disease and Its Associations with Clinical Characteristics**
Parkinson Hastalığında Sarkopeni Risk Algısı ve Klinik Özelliklerle İlişkisi
Hatice Betigül Meral, Abdülkadir Ermiş, Aygül Tantik Pak, Mehmet Yalman, Eren Toplutaş, Ece Akın, Sena Tolu; İstanbul, Turkey

414 **Analysis of Patients Who Underwent Surgery for Gastric Gastrointestinal Stromal Tumors**
Mide Gastrointestinal Stromal Tümör Nedeniyle Opere Edilen Hastaların Analizi
Nurettin Şahin, Turgut Dönmez, Mahmut Said Değerli, Ferman Tevfik Özyalvaç, Süleyman Büyükaşık, Allanur Mametkulyyev, Duhan Kanmaz, Ahmet Sürek; İstanbul, Turkey

CASE REPORT / OLGU SUNUMU

427 **Occult Bronchial Carcinoid Tumor: A Rare Case of Lung Cancer**
Okült Bronşiyal Karsinoid Tümör: Nadir Bir Akciğer Kanseri Olgusu
Hülya Abalı, Ayşenur İlkyaz, Çiğdem Sabancı, Sida Gösterici, Fatma Tokgöz Akyıl; İstanbul, Turkey

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):279-289
DOI: 10.4274/BMB.galenos.2026.81994

Retrospective Analysis of Vitamin B12 and Folic Acid Levels Before and After Antiepileptic (Anticonvulsant) Treatment in Patients with Neuropathic Pain

Nöropatik Ağrı Nedeniyle Antiepileptik (Antikonvülzan) Tedavi Alan Hastalarda B12 ve Folik Asit (Folat) Düzeyinin Tedavi Öncesi ve Sonrası Retrospektif İncelenmesi

Deniz Öke, Zeynel Karakullukçuoğlu

University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Turkey

Abstract

Objective: The aim of this study was to investigate the effects of commonly prescribed antiepileptic drugs (pregabalin and gabapentin) on serum vitamin B12 and folate levels in patients with neuropathic pain (NP).

Method: This retrospective study was conducted by reviewing the electronic records of 96 adult patients diagnosed with NP who received treatment with pregabalin or gabapentin between 1 June 2024 and 31 December 2024. Ethics committee approval was obtained on 8 January 2025, and data access and statistical analyses were initiated after this date. Serum vitamin B12, folate, and hematological parameters were obtained from routine laboratory tests performed before antiepileptic drug initiation and from the first available follow-up tests conducted at least three months after treatment initiation. Statistical analyses included the Wilcoxon signed-rank test and paired-samples t-test, as appropriate for within-patient comparisons.

Results: Of 287 potentially eligible patients screened, 96 met the eligibility criteria and were included in the final analysis. Compared with baseline values, post-treatment vitamin B12 levels were lower (from 416.0 pg/mL to 282.0 pg/mL, p=0.001), and folate levels also decreased (from 8.15 ng/mL to 7.45 ng/mL, p=0.003). In subgroup analyses, patients receiving pregabalin showed statistically significant

Öz

Amaç: Bu çalışmanın amacı, nöropatik ağrı (NP) tedavisinde sıklıkla kullanılan antiepileptik ilaçların (pregabalin ve gabapentin) serum vitamin B12 ve folat düzeyleri üzerindeki etkilerini araştırmaktır.

Yöntem: Bu retrospektif çalışma, 1 Haziran 2024 ile 31 Aralık 2024 tarihleri arasında pregabalin veya gabapentin tedavisi gören ve NP tanısı konmuş 96 yetişkin hastanın elektronik kayıtlarının incelenmesiyle gerçekleştirilmiştir. Etik kurul onayı 8 Ocak 2025 tarihinde alınmış olup, veri erişimi ve istatistiksel analizler bu tarihten sonra başlatılmıştır. Serum B12 vitamini, folat ve hematolojik parametreler, antiepileptik ilaç tedavisi başlamadan önce yapılan rutin laboratuvar testlerinden ve tedavi başlangıcından en az üç ay sonra gerçekleştirilen ilk takip testlerinden elde edildi. İstatistiksel analizlerde, hasta içi eşleştirilmiş karşılaştırmalar için veri dağılımına göre Wilcoxon işaretli sıralar testi veya eşleştirilmiş örneklem t-testi kullanıldı.

Bulgular: Potansiyel olarak uygun görülen 287 hastadan 96'sı uygunluk kriterlerini karşıladı ve nihai analize dahil edildi. Tedavi öncesi değerlerle karşılaştırıldığında, tedavi sonrası B12 vitamini düzeyleri daha düşüktü (416,0 pg/mL'den 282,0 pg/mL'ye, p=0,001) ve folat düzeylerinde de azalma gözlendi (8,15 ng/mL'den 7,45 ng/mL'ye, p=0,003). Alt grup analizlerinde, pregabalin alan hastalarda

Address for Correspondence: Deniz Öke, University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, Department of Physical Medicine and Rehabilitation, İstanbul, Turkey
E-mail: dr_denizoke@hotmail.com **ORCID:** orcid.org/0000-0003-3213-0363
Received: 30.09.2025 **Accepted:** 16.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Öke D, Karakullukçuoğlu Z. Retrospective analysis of vitamin B12 and folic acid levels before and after antiepileptic (anticonvulsant) treatment in patients with neuropathic pain. Bagcilar Med Bull. 2026;11(3):279-289

Abstract

decreases in both vitamin B12 (from 448.0 pg/mL to 292.5 pg/mL, p=0.001) and folate levels (from 7.9 ng/mL to 7.5 ng/mL, p=0.007), while in those treated with gabapentin, significant reductions were observed in vitamin B12 (from 368.0 pg/mL to 273.0 pg/mL, p=0.001) and vitamin D levels (from 23.9 ng/mL to 15.7 ng/mL, p=0.021). Sex-stratified analyses showed significant within-group reductions in vitamin B12 and folate levels in both male and female patients.

Conclusion: Pregabalin and gabapentin, widely used in the treatment of NP, are associated with significant reductions in serum vitamin B12 and folate levels. These findings suggest a potential association between antiepileptic drug use and vitamin level reduction. Routine vitamin monitoring may be considered in clinical practice; however, these observational findings do not establish causality, and further prospective studies are required to confirm these associations.

Keywords: Antiepileptic drugs, folate, gabapentin, neuropathic pain, pregabalin, vitamin B12

Öz

hem B12 vitamini (448,0 pg/mL'den 292,5 pg/mL'ye, p=0,001) hem de folat (7,9 ng/mL'den 7,5 ng/mL'e, p=0,007) düzeylerinde istatistiksel olarak anlamlı düşüşler görülürken, gabapentin ile tedavi edilenlerde B12 vitamini (368,0 pg/mL'den 273,0 pg/mL'ye, p=0,001) ve D vitamini (23,9 ng/mL'den 15,7 ng/mL'ye, p=0,021) düzeylerinde anlamlı düşüşler gözlendi. Cinsiyete göre tabakalandırılmış analizlerde, hem erkeklerde hem de kadınlarda B12 vitamini ve folat düzeylerinde grup içi anlamlı azalmalar saptandı.

Sonuç: NP tedavisinde yaygın olarak kullanılan pregabalin ve gabapentin, serum vitamin B12 ve folat düzeylerinde anlamlı azalma ile ilişkili bulunmuştur. Bu bulgular, antiepileptik ilaç kullanımı ile vitamin seviyesindeki azalma arasında potansiyel bir ilişki-olduğunu düşündürmektedir. Klinik uygulamada rutin vitamin takibi düşünülebilir; ancak bu gözlemsel bulgular nedensellik ilişkisini kanıtlamamaktadır ve bu bağlantıları doğrulamak için daha fazla prospektif çalışma gereklidir.

Anahtar kelimeler: Antiepileptik ilaçlar, folat, gabapentin, nöropatik ağrı, pregabalin, vitamin B12

Introduction

Neuropathic pain (NP) is a chronic condition that arises as a direct consequence of a lesion or disease affecting the somatosensory nervous system. It is characterized by maladaptive changes in neural processing that lead to dysfunctional pain signaling pathways and abnormal excitability of peripheral and central neurons (1-4). With a global prevalence of 2-8%, NP significantly impairs quality of life, reduces functional capacity, and imposes considerable socioeconomic costs on patients and healthcare systems (5-7). Despite the availability of multiple therapeutic options, a large proportion of patients remain undertreated or receive suboptimal management (8,9).

The pathophysiology of NP is highly complex and involves peripheral and central sensitization, maladaptive neuroplasticity, and alterations in neurotransmitter signaling (10,11). These processes contribute to heterogeneous clinical presentations and variable treatment responses, which complicates effective management (4). Current guidelines, including the French recommendations established in 2020 based on the GRADE system, emphasize a stepwise pharmacological approach. First-line therapies include serotonin-norepinephrine reuptake inhibitors such as duloxetine and venlafaxine, tricyclic antidepressants, and gabapentinoids (gabapentin and pregabalin), in addition to topical lidocaine and non-pharmacological methods such as transcutaneous electrical nerve stimulation for peripheral NP (12-14). Second-line strategies often involve combination therapies, botulinum toxin A, or high-concentration capsaicin patches, while third-line approaches include spinal cord stimulation, repetitive transcranial magnetic stimulation, and strong opioids, typically reserved for refractory cases (15,16).

Antiepileptic drugs (AEDs), particularly pregabalin and gabapentin, play a crucial role in NP treatment by modulating excitatory neurotransmission through calcium channel blockade (17,18). However, there is growing evidence that chronic AED therapy may negatively affect nutritional status, particularly by reducing serum vitamin B12 and folate levels (19,20). Vitamin B12 deficiency is known to impair myelin synthesis and nerve conduction, potentially worsening neuropathic symptoms or contributing to cognitive decline (19,20). Similarly, folate is essential for DNA synthesis, repair, and methylation, as well as for neuronal health and regeneration (20). Deficiencies in either vitamin may lead to elevated homocysteine levels, which are associated with neurotoxicity and increased cardiovascular risk (1,21,22).

Gabapentinoids are widely used in NP, and growing evidence suggests that AEDs may influence vitamin metabolism. Therefore, evaluating changes in vitamin B12 and folate levels in this population is clinically important. Monitoring these vitamins before and after therapy may help optimize treatment strategies and prevent secondary complications. Accordingly, this study aimed to investigate the association between AED therapy and changes in serum vitamin B12 and folate levels in patients with NP.

Materials and Methods

Study Design and Participants

This study was designed as a retrospective observational analysis. Electronic medical records of adult patients with NP who received pregabalin or gabapentin between 1 June 2024 and 31 December 2024 were reviewed. Eligible participants were identified through the hospital information system based on diagnosis codes and prescription records. Ethics committee approval was obtained on 8 January 2025 (University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, no: 85, date: 08.01.2025), and data extraction as well as statistical analysis were initiated only after this approval, in accordance with national regulations for retrospective data use. As this was a retrospective study based on previously recorded clinical data, the requirement for obtaining informed consent was waived by the Ethics Committee in accordance with institutional policy and national regulations. The study used anonymized data from routine clinical practice, with no direct patient contact or additional tests, and was conducted in accordance with national regulations governing retrospective chart reviews. No direct patient contact occurred, and no identifiable personal information was used. The study had no prospective interventional component; all data were obtained from pre-existing medical records without any additional visits, laboratory tests, or treatment modifications beyond routine clinical care. Eligible participants were identified through electronic medical records. Inclusion criteria required patients to be between 20 and 75 years of age, to have an NP diagnosis confirmed by the DN-4 questionnaire, and to be receiving AED therapy with pregabalin or gabapentin. Additionally, participants had to have serum vitamin B12 and folate levels measured both before and after AED treatment, with no history of vitamin B12 or folate supplementation during therapy. Patients with comorbid conditions that could potentially influence vitamin status, including diabetes mellitus, liver failure, or chronic kidney disease, were excluded from the study.

Patients who met these criteria and had completed at least six months of AED therapy were included in the analysis. Exclusion criteria were: Presence of systemic diseases (e.g., diabetes mellitus, hepatic or renal insufficiency), prior or concurrent vitamin supplementation, and missing laboratory measurements at baseline or follow-up.

Because several medical conditions and commonly used medications may independently influence serum vitamin B12 and folate concentrations, potential confounders were carefully reviewed through electronic medical records. Patients receiving medications known to reduce vitamin B12 or folate levels—such as metformin, proton pump inhibitors, H2-blockers, or long-term antifolate agents—were excluded from the study. Additionally, patients with clinically significant hypothyroidism, untreated hyperlipidemia, or other systemic conditions that could alter vitamin absorption or metabolism were excluded unless laboratory data confirmed stable disease and adequate control prior to initiation of antiepileptic therapy. These measures were implemented to minimize confounding and improve internal validity.

Data Collection and Laboratory Measurements

Patient data, including demographic information (age, sex), comorbidities, type of AED used (pregabalin or gabapentin), daily drug dose, and duration of therapy, were retrospectively extracted from hospital records.

Biochemical analyses included serum vitamin B12, folate, and vitamin D levels, measured at two time points: (1) At baseline, defined as the laboratory assessment performed within 0-4 weeks before AED initiation, and (2) at the first available follow-up measurement obtained at least 3 months after treatment initiation. Because of the retrospective design, the timing of follow-up tests was not protocol-driven but reflected routine clinical practice; therefore, follow-up intervals varied between patients and were not standardized. The reported median treatment duration of 16.5 months refers to the total duration of AED use at the time of data extraction and does not represent the fixed interval between baseline and follow-up laboratory measurements. Hematological parameters such as hemoglobin, hematocrit, mean corpuscular volume (MCV), red blood cell count, platelet count, mean platelet volume, plateletcrit (PCT), white blood cell (WBC) count, and differential leukocyte counts were also recorded when available. For each laboratory parameter, paired analyses were restricted to patients with both baseline and post-treatment measurements; sex- and drug-specific subgroup analyses were conducted within the corresponding analyte-specific paired cohort.

The reference ranges applied in this study were 180-920 pg/mL for vitamin B12, 3-17 ng/mL for folic acid (folate), and 20-50 ng/mL for vitamin D. These intervals reflect widely accepted clinical laboratory standards used to evaluate vitamin status in adults. The chosen cut-off values are consistent with commonly reported ranges in the literature,

where serum vitamin B12 concentrations below 200 pg/mL indicate deficiency and values between 200-300 pg/mL suggest borderline status (23,24). Similarly, folate levels below 3 ng/mL are typically considered deficient, while levels above 4 ng/mL are regarded as adequate for normal hematopoietic and neurological function (1,25). For vitamin D, concentrations below 20 ng/mL are generally defined as deficient, whereas levels between 20-50 ng/mL are considered sufficient for bone and musculoskeletal health (26,27). These standardized thresholds ensure comparability with prior studies and allow for reliable interpretation of laboratory findings in the context of NP and AED therapy.

Serum vitamin D concentrations were measured using a standardized chemiluminescence immunoassay. Vitamin B12 and folate levels were analyzed using immunoassay-based methods, and all reference intervals corresponded to manufacturer-provided ranges validated by the hospital's central biochemistry laboratory. All assays were performed according to the manufacturer's instructions, following routine internal quality control protocols.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics version 27.0 (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as median (minimum-maximum) values, and categorical variables were presented as frequencies and percentages. The normality of continuous variables was assessed using the Shapiro-Wilk test. For paired comparisons of baseline and post-treatment values, the Wilcoxon signed-rank test was used for non-normally distributed variables, whereas the paired-sample t-test was applied when normality assumptions were met. For paired comparisons of baseline and post-treatment values in the overall cohort and within sex- and drug-specific subgroups, the Wilcoxon signed-rank test was used for non-normally distributed variables, whereas the paired-samples t-test was applied when the normality assumption was met. The p-values reported for the sex-stratified analyses represent within-sex paired comparisons. No formal between-sex comparison of individual change scores was performed; therefore, numerical differences between male and female strata were not interpreted as evidence of sex-specific treatment effects. In addition to p-values, effect sizes for paired non-parametric comparisons were calculated using the rank-biserial correlation (rrb). For descriptive purposes, the percentage change in group medians was calculated for vitamin B12 and folate as [(post-treatment group median–baseline group median)/baseline group median] ×100. Negative values indicate a reduction. These values were calculated from the group medians and do not represent the median of individual patient-level percentage changes; they were not used for inferential testing. Because the primary outcomes were predefined, no formal multiple-comparison correction was applied; however, secondary analyses were interpreted cautiously. A p-value <0.05 was considered statistically significant. Because this was an observational retrospective study, no a priori sample size calculation was performed. Subgroup analyses were exploratory in nature, and statistical power may be reduced in smaller strata, particularly in gender- and drug-specific comparisons. The analyses were primarily designed to characterize within-patient changes in vitamin levels. Multivariable regression was not prespecified, and the available sample size was considered modest for simultaneous adjustment for multiple potential confounders. Therefore, the subgroup analyses were considered exploratory, and the observed associations were interpreted cautiously without attributing the changes independently to drug exposure.

Results

During the study period, a total of 287 patients with neuropathic pain who had been prescribed pregabalin or gabapentin between 1 June 2024 and 31 December 2024 were screened for eligibility through the electronic medical record system. Of these, 191 patients were excluded according to the predefined eligibility criteria: 64 because of missing baseline vitamin B12 and/or folate data, 58 because of missing post-treatment vitamin B12 and/or folate data, 23 because they received vitamin B12 or folate supplementation during the treatment period, 27 because of comorbid conditions potentially affecting vitamin B12 or folate metabolism, 14 because of concomitant medications known to interfere with vitamin B12 or folate status, and 5 because they did not meet other eligibility criteria. After these exclusions, 96 patients fulfilled all eligibility criteria and were included in the final analysis, of whom 59 received pregabalin and 37 received gabapentin (Figure 1).

Demographic and clinical characteristics of patients are presented in Table 1. The median age of the included patients was 61.0 years (range, 20-75 years). The cohort consisted of 67 males (69.8%) and 29 females (30.2%). Regarding comorbidities, 30.2% had no comorbidity, while 21.9% had multiple comorbid conditions. The most common individual comorbidities were hypertension (13.5%), hypothyroidism (10.4%), and hyperlipidemia (6.3%). In terms of medication, 61.5% of the patients were using pregabalin, and 38.5% were using gabapentin. The

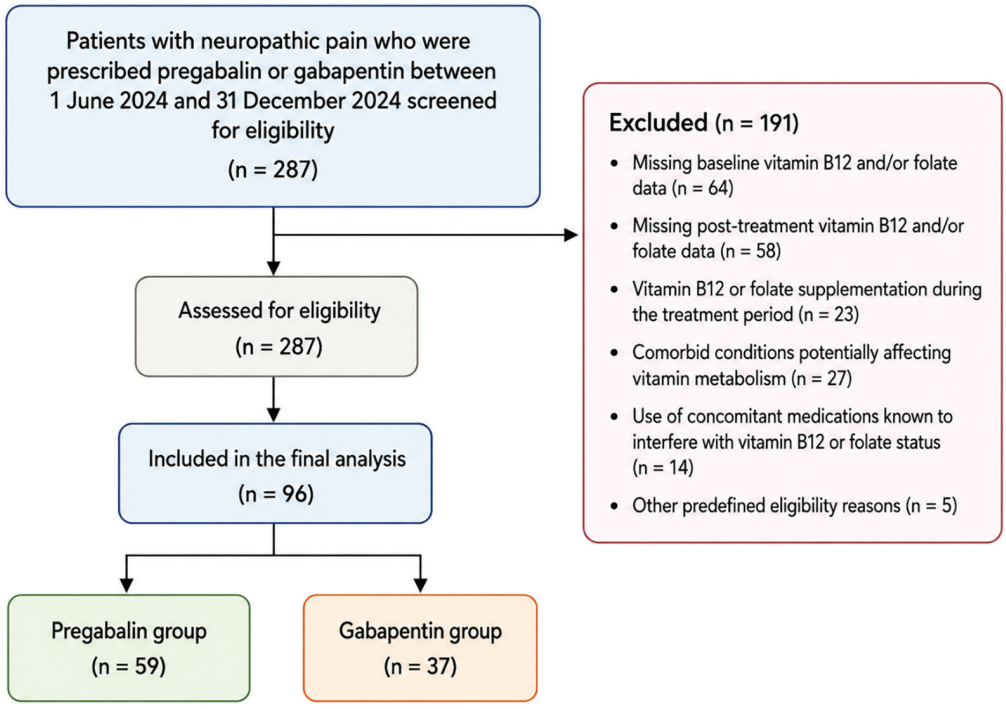

Figure 1. Flow diagram of patient screening, exclusion, and inclusion in the final study population

median drug dose was 300.00 mg/day (25.00-2400.00), and the duration of use was 16.5 months (3.00-149.00). This duration reflects the total exposure to AED therapy at the time of data extraction. Baseline and post-treatment laboratory results demonstrated statistically significant changes in vitamin levels. Other hematological parameters included a median hemoglobin of 13.20 g/dL, hematocrit of 39.80%, MCV of 83.30 fL, and platelet count of 262.00×10^3/µL. Inflammatory and immune cell markers such as WBC, neutrophils, lymphocytes, and others were within expected ranges (Table 1).

The comparison of baseline and post-treatment laboratory values demonstrated statistically significant changes in vitamin B12 and folic acid levels during AED treatment; however, these changes should be interpreted as associations rather than proof of causality. The median vitamin B12 level significantly decreased from 416.00 pg/mL (range: 145.00-1993.00) at baseline to 282.00 pg/mL (117.00-1216.00) after treatment (p=0.001). The median folic acid level declined from 8.15 ng/mL (2.50-22.00) to 7.45 ng/mL (3.40–17.10), also reaching statistical significance (p=0.003). In contrast, the change in vitamin D levels from 20.55 ng/mL (5.00-51.90) at baseline to 19.05 ng/mL (5.70-52.30) after treatment did not reach statistical significance (p=0.122). The percentage changes in group medians were -32.2% for vitamin B12 and -8.6% for folate. The effect size (rank-biserial correlation) indicated a large effect for the decline in vitamin B12 (rrb=0.62) and a moderate effect for the reduction in folate (rrb=0.34) (Table 2). Sex-stratified within-group analyses showed statistically significant reductions in vitamin B12 and folate levels from baseline to post-treatment in both male and female patients. Among males, median vitamin B12 decreased from 430.00 pg/mL (145.00-1757.00) to 273.00 pg/mL (137.00-1216.00), and folic acid from 8.35 ng/mL (2.50-22.00) to 7.60 ng/mL (3.40-17.10), with p-values of 0.001 and 0.028, respectively. A significant reduction was also observed in vitamin D levels among males (p=0.040). In female patients, vitamin B12 levels declined significantly from 407.50 pg/mL (169.00-1993.00) to 311.50 pg/mL (117.00-699.00) (p=0.001), and folic acid dropped from 8.00 ng/mL (3.20-15.00) to 6.60 ng/mL (3.90-14.60) (p=0.038). However, the change in vitamin D levels among female patients, from 19.00 ng/mL (5.00-35.00) at baseline to 20.80 ng/mL (8.00-52.30) after treatment, was not statistically significant (p=0.066). Effect sizes for gender-specific analyses demonstrated a large effect for vitamin B12 decline in males (rrb=0.59) and a moderate effect in females (rrb=0.41) (Table 3). In the comparison of baseline and post-treatment hematological parameters, significant decreases were observed in several

Table 1. Baseline demographic, clinical, and hematological characteristics of the entire study cohort

Parameter (unit)	Median (min-max) or n (%)
Age (years)	61.00 (20.00-75.00)
Gender-male	67 (69.8%)
Gender-female	29 (30.2%)
Comorbidity	
No comorbidity	29 (30.2%)
Hypertension	13 (13.5%)
Hypothyroidism	10 (10.4%)
Hyperlipidemia	6 (6.3%)
Respiratory disease	3 (3.1%)
Cardiac disease	5 (5.2%)
Multiple	21 (21.9%)
Other	8 (8.3%)
Drug type	
Pregabalin	59 (61.5%)
Gabapentin	37 (38.5%)
Drug dose (mg/day)	300.00 (25.00-2400.00)
Duration of use (months)	16.50 (3.00-149.00)
RBC (10^6/µL)	4.78 (3.44-6.16)
Hemoglobin (g/dL)	13.20 (7.90-15.70)
Hematocrit (%)	39.80 (30.70-48.90)
MCV (fL)	83.30 (67.00-100.20)
MCH (pg)	27.20 (21.80-32.40)
MCHC (g/dL)	32.70 (28.10-36.00)
RDW (%)	13.90 (10.90-23.40)
Platelets (10^3/µL)	262.00 (132.00-494.00)
MPV (fL)	9.30 (6.60-13.90)
PDW (%)	13.80 (10.00-20.00)
PCT (%)	0.23 (0.11-0.47)
WBC (10^3/µL)	7.13 (3.10-15.20)
Neutrophils (10^3/µL)	4.40 (1.30-12.30)
Lymphocytes (10^3/µL)	2.00 (0.60-4.40)
Monocytes (10^3/µL)	0.49 (0.18-1.20)
Eosinophils (10^3/µL)	0.12 (0.01-0.66)
Basophils (10^3/µL)	0.02 (0.00-0.27)

RBC: Red blood cell count, MPV: Mean platelet volume, PCT: Plateletcrit, WBC: White blood cell, RDW: Red cell distribution width, MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration

indices. Platelet counts showed a notable reduction from a median of 264.0×10^3/µL to 239.0×10^3/µL (p=0.033). PCT levels declined significantly from 0.3% to 0.2% (p=0.008). Among leukocyte parameters, the total WBC count decreased from 7250.0/µL to 7005.0/µL (p=0.031), while neutrophil counts declined from 4300.0/µL to 3880.0/µL (p=0.043). No statistically significant changes were detected in red blood cell indices, hemoglobin, hematocrit, lymphocyte counts, or platelet volume indices (p>0.05).

In patients treated with pregabalin, a significant reduction in vitamin B12 levels was observed, with median values decreasing from 448.0 pg/mL at baseline to 292.5 pg/mL post-treatment (p=0.001). Similarly, serum folate concentrations declined from 7.9 ng/mL to 7.5 ng/mL, reaching statistical significance (p=0.007). Hematological analysis revealed a significant decrease in platelet counts, from 283.0×10^3/µL to 240.5×10^3/µL (p=0.037), as well as a reduction in PCT levels from 0.3% to 0.2% (p=0.013). In contrast, no significant changes were noted in vitamin D levels, red blood cell indices, hemoglobin, hematocrit, or leukocyte counts (p>0.05) (Table 4). In patients receiving gabapentin therapy, serum vitamin B12 levels showed a significant decline, with median values decreasing from 368.0 pg/mL at baseline to 273.0 pg/mL after treatment (p=0.001). Vitamin D concentrations also dropped significantly, from 23.9 ng/mL to 15.7 ng/mL (p=0.021). In addition, MCHC demonstrated a slight but statistically significant reduction, from 33.1 g/dL to 32.5 g/dL (p=0.047). No significant changes were observed in folate levels, red blood cell indices, platelet parameters, or leukocyte counts (p>0.05) (Table 5).

Discussion

This study provides observational evidence regarding the association between AED therapy and alterations in vitamin B12 and folate status in patients with NP. Our findings demonstrate statistically significant reductions in both vitamins over the course of AED treatment; however, due to the retrospective design, these changes cannot be interpreted as causal effects.

Specifically, the group median folate level declined from 8.15 ng/mL to 7.45 ng/mL (p=0.003), while the group median vitamin B12 level decreased from 416.0 pg/mL to 282.0 pg/mL (p=0.001). These values correspond to percentage changes in group medians of -8.6% for folate and -32.2% for vitamin B12. Consistent with our findings, Huang et al. (20) also reported significant decreases in serum folate and vitamin B12 levels among epilepsy patients following AED administration. Similar reductions reported in previous studies (20) support the possibility that AEDs may influence vitamin metabolism, but prospective and mechanistic studies are needed to confirm this hypothesis.

The clinical implications of these biochemical changes are noteworthy. Vitamin B12 deficiency is known to cause

Table 2. Comparison of baseline and post-treatment laboratory parameters

Parameter (unit)	Baseline	Post-treatment	p-value
Vitamin B12 (pg/mL)	416.00 (145.00-1993.00)	282.00 (117.00-1216.00)	**0.001**
Folic acid (ng/mL)	8.15 (2.50-22.00)	7.45 (3.40-17.10)	**0.003**
Vitamin D (ng/mL)	20.55 (5.00-51.90)	19.05 (5.70-52.30)	0.122

Data are presented as median (minimum-maximum). For each parameter, paired analyses included patients with both baseline and post-treatment measurements

Table 3. Paired comparison of baseline and post-treatment laboratory parameters stratified by sex

Gender	Parameter (unit)	Baseline	Post-treatment	p-value
Male	Vitamin B12 (pg/mL)	430.00 (145.00-1757.00)	273.00 (137.00-1216.00)	**0.001**
Male	Folic acid (ng/mL)	8.35 (2.50-22.00)	7.60 (3.40-17.10)	**0.028**
Male	Vitamin D (ng/mL)	20.90 (6.50-51.90)	17.90 (5.70-41.00)	**0.040**
Female	Vitamin B12 (pg/mL)	407.50 (16.00-1993.00)	311.50 (117.00-699.00)	**0.001**
Female	Folic acid (ng/mL)	8.00 (3.20-15.00)	6.60 (3.90-14.60)	**0.038**
Female	Vitamin D (ng/mL)	19.00 (5.0-35.0)	20.80 (8.0-50.60)	0.066

Data are presented as median (minimum-maximum). The reported p-values represent paired baseline-to-post-treatment comparisons performed separately within male and female patients. No direct between-sex comparison of individual changes was performed

Table 4. Comparison of baseline and post-treatment parameters in patients receiving pregabalin

Parameter (unit)	Baseline	Post-treatment	p-value
Vitamin B12 (pg/mL)	448.0 (145.0-1993.0)	292.5 (147.0-770.0)	**0.001**
Vitamin D (ng/mL)	19.4 (5.00-51.90)	20.9 (5.70-52.30)	0.904
Folate (ng/mL)	7.9 (2.5-22.0)	7.5 (3.9-17.1)	**0.007**
RBC ($\times 10^6$/µL)	4.5 (3.5-5.5)	4.6 (3.4-5.8)	0.265
Hemoglobin (g/dL)	12.5 (7.9-15.7)	12.9 (9.5-15.7)	0.394
Hematocrit (%)	37.8 (27.9-46.1)	38.8 (31.8-46.6)	0.521
MCV (fL)	86.5 (60.5-100.7)	85.8 (66.9-99.5)	0.739
MCH (pg)	28.4 (17.2-32.7)	28.0 (21.3-33.7)	0.975
MCHC (g/dL)	32.5 (28.5-35.1)	32.6 (29.8-34.8)	0.322
Platelets ($\times 10^3$/µL)	283.0 (176.0-411.0)	240.5 (56.0-442.0)	**0.037**
MPV (fL)	9.9 (1.9-12.7)	9.6 (7.6-13.9)	0.804
PDW (%)	16.0 (11.4-22.8)	15.9 (10.5-17.8)	0.604
PCT (%)	0.3 (0.2-0.4)	0.2 (0.1-0.4)	**0.013**
WBC-total ($\times 10^3$/µL)	7.35 (3.85-13.63)	7.45 (2.89-10.03)	0.143
Neutrophils ($\times 10^3$/µL)	4.34 (1.17-9.46)	4.15 (1.80-6.79)	0.222
Lymphocytes($\times 10^3$/µL)	2.45 (1.06-4.24)	2.35 (0.34-3.55)	0.528

RBC: Red blood cell count, MPV: Mean platelet volume, PCT: Plateletcrit, WBC: White blood cell, RDW: Red cell distribution width, MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration

peripheral neuropathy, raising the possibility of a paradox in which medications used to alleviate NP may simultaneously exacerbate underlying neuronal dysfunction (28). Indeed, prior studies reported subnormal vitamin B12 levels in approximately 44.8% of patients treated with carbamazepine, suggesting that gabapentinoids may induce comparable metabolic effects (19,29).

Several mechanisms may contribute to AED-related vitamin deficiencies. These include impaired vitamin B12 absorption due to increased intestinal pH, inhibition of folate-metabolizing enzymes, and competitive interference between folate and AEDs at intestinal absorption sites (1,17,19,22). Previous evidence also indicates that long-term AED therapy alters folate and vitamin B12 bioavailability, particularly with phenytoin. Moreover, calcium channel-acting drugs such as pregabalin and gabapentin may inhibit transport mechanisms required for vitamin uptake (30,31).

Sex-stratified analyses showed significant within-group reductions in vitamin B12 and folate levels among both male and female patients. However, individual changes

Table 5. Comparison of baseline and post-treatment parameters in patients receiving gabapentin

Parameter (unit)	Baseline	Post-treatment	p-value
Vitamin B12 (pg/mL)	368.0 (169.0-1000.0)	273.0 (117.0-1216.0)	**0.001**
Vitamin D (ng/mL)	23.9 (7.1-50.6)	15.7 (6.0-33.7)	**0.021**
Folate (ng/mL)	8.9 (3.2-15.6)	7.5 (3.4-14.6)	0.199
RBC ($\times10^6$/µL)	4.5 (4.0-6.8)	4.5 (3.9-6.0)	0.931
Hemoglobin (g/dL)	13.3 (9.4-14.8)	13.4 (9.9-15.6)	0.509
Hematocrit (%)	39.6 (30.9-44.3)	40.8 (31.9-48.1)	0.569
MCV (fL)	88.5 (63.1-96.4)	86.5 (58.4-100.3)	0.280
MCH (pg)	28.8 (20.4-32.7)	28.9 (19.0-32.8)	0.279
MCHC (g/dL)	33.1 (30.2-34.8)	32.5 (31.0-34.4)	**0.047**
Platelets ($\times10^3$/µL)	250.0 (130.0-447.0)	233.0 (128.0-437.0)	0.345
MPV (fL)	9.9 (7.1-13.3)	9.7 (7.8-13.9)	0.959
PDW (%)	16.2 (9.4-16.7)	16.1 (15.0-16.7)	0.988
PCT (%)	0.2 (0.1-0.3)	0.2 (0.1-0.4)	0.352
WBC-total ($\times10^3$/µL)	7.09 (4.29-10.99)	6.68 (4.82-9.47)	0.088
Neutrophils ($\times10^3$/µL)	4.23 (2.29-7.56)	3.71 (2.31-7.10)	0.064
Lymphocytes($\times10^3$/µL)	2.14 (1.19-4.25)	2.16 (1.10-3.03)	0.584

RBC: Red blood cell count, MPV: Mean platelet volume, PCT: Plateletcrit, WBC: White blood cell, RDW: Red cell distribution width, MCV: Mean corpuscular volume, MCH: Mean corpuscular hemoglobin, MCHC: Mean corpuscular hemoglobin concentration

were not formally compared between sexes. Therefore, the observed numerical differences in percentage change should not be interpreted as evidence of sex-specific effects or used to support sex-specific supplementation strategies. Future studies should evaluate whether sex modifies these changes through direct comparisons of individual change scores or longitudinal models incorporating a time-by-sex interaction.

Importantly, the present findings do not establish that the observed reductions in vitamin levels are independently attributable to pregabalin or gabapentin. Age and sex may influence vitamin status and therefore remain potential confounders in this retrospective cohort. Although major medical conditions and medications known to interfere with vitamin B12 or folate metabolism were excluded, residual confounding related to age, sex, baseline vitamin status, dietary factors, treatment duration, drug dose, and other unmeasured characteristics cannot be excluded. Accordingly, the observed pre-post differences should be interpreted as associations rather than drug-specific causal effects. Larger prospective studies incorporating multivariable models with changes in vitamin B12 and folate as outcomes and age, sex, drug type, dose, treatment duration, and baseline vitamin concentrations as covariates are needed to determine whether gabapentinoid exposure is independently associated with vitamin depletion.

Given the observed associations, closer monitoring of vitamin B12 and folate levels may be considered in selected high-risk patients undergoing AED therapy for NP; however, routine supplementation or systematic screening cannot be recommended solely on the basis of this single retrospective study. Future prospective and randomized controlled trials are required to determine whether vitamin monitoring and supplementation improve clinical outcomes in NP. Previous reports have described variable vitamin B12 outcomes under long-term AED use, ranging from decreases to normal or even elevated levels. Importantly, deficiencies may arise weeks to months after initiation of therapy (32). Therefore, these findings may suggest a potential role for closer monitoring and individualized supplementation strategies, particularly in high-risk groups such as elderly individuals, malnourished patients, or those with multiple comorbidities (33). Of note, a recent retrospective cohort study reported that repeated gabapentin prescriptions among adults with chronic low back pain were associated with increased risks of dementia and mild cognitive impairment, particularly among adults aged 18-64 years (34). Taken together, these observations highlight the potential clinical relevance of monitoring vitamin status; however, further prospective studies are required before establishing definitive clinical recommendations.

From a health economics perspective, routine supplementation may represent a potentially cost-effective strategy; however, this requires confirmation in future studies. Vitamin B12 and folic acid insufficiency contribute to elevated homocysteine levels, which have been implicated as independent risk factors for cardiovascular, cerebrovascular, and peripheral vascular diseases (35,36). Hyperhomocysteinemia-associated peripheral neuropathy, in combination with vascular complications, substantially increases healthcare costs and patient morbidity (37). Timely supplementation may therefore prevent such complications while simultaneously enhancing treatment outcomes and quality of life.

Study Limitations

This study has some limitations that should be acknowledged. First, the relatively short follow-up period may not adequately reflect long-term alterations in vitamin status associated with AED therapy. Second, functional outcomes and quality-of-life measures were not assessed, which limits the clinical interpretation of the observed biochemical changes. Previous studies have reported a dose-dependent relationship between certain AEDs, including pregabalin and topiramate, and serum vitamin B12 and folate concentrations; however, our study was not designed to evaluate such associations in detail. Additional limitations include the relatively small sample size, the retrospective design, and the use of two different AEDs with varying dosing regimens, which may have introduced heterogeneity and limited the ability to establish a clear dose–response relationship. Although major confounding conditions were screened and patients with uncontrolled systemic diseases or medications known to interfere with vitamin metabolism were excluded, the retrospective design does not fully eliminate the possibility of residual confounding. Furthermore, multivariable regression adjusting simultaneously for age, sex, drug type, drug dose, treatment duration, and baseline vitamin concentrations was not performed. Therefore, the present study cannot determine the independent contribution of gabapentinoid therapy to the observed changes in vitamin levels, and residual confounding remains possible. Chronic illnesses such as hypothyroidism or hyperlipidemia, even when stable, may have subtle effects on vitamin homeostasis, and these factors could contribute to variability in the observed results. Another limitation is the modest sample size, particularly after stratification by drug type and gender. Smaller subgroup sizes reduce statistical power and may limit the robustness of the observed effect sizes and subgroup-specific findings. Therefore, these results should be interpreted with caution, and larger prospective studies are warranted to validate these subgroup differences.

Future randomized controlled trials are warranted to assess the effects of systematic vitamin supplementation on rehabilitation outcomes, functional capacity, and quality of life in patients with NP. Such studies should also determine optimal AED dosing strategies, appropriate timing for supplementation, and identify high-risk patient subgroups. Addressing these aspects would provide more robust evidence to guide clinical practice and inform the development of comprehensive treatment guidelines.

Implications for Future Research

Future research should focus on prospective, large-scale randomized controlled trials to further elucidate the relationship between gabapentinoid therapy and alterations in vitamin B12 and folate levels. Such studies should not only evaluate long-term biochemical changes but also incorporate functional and clinical outcomes, including neuropathic symptom progression, cognitive performance, and quality of life. Investigating potential dose–response relationships, differences between various AED classes, and the impact of treatment duration will provide a more comprehensive understanding of the mechanisms underlying vitamin depletion. Furthermore, studies should aim to identify high-risk subgroups, such as elderly patients, those with nutritional deficiencies, or individuals with multiple comorbidities, to inform targeted supplementation strategies. Incorporating cost-effectiveness analyses and health economic perspectives would also be valuable in incorporating cost-effectiveness analyses and health economic perspectives would also be valuable in shaping evidence-based guidelines for vitamin monitoring in clinical practice; however, further prospective studies are required.

Conclusion

In conclusion, this study demonstrates that AED therapy use is associated with significant reductions in serum vitamin B12 and/or folate levels in patients with NP. These findings suggest a potential relationship between antiepileptic drug use and altered vitamin status. Routine monitoring of vitamin levels may be considered in clinical practice, particularly in high-risk populations; however, further prospective studies are required to confirm these findings and to establish clinical recommendations.

The results further suggest that integrating nutritional assessment and vitamin supplementation into NP management may be beneficial; however, the potential clinical impact of such approaches should be interpreted with caution and requires validation in prospective studies. Such an approach may be consistent with multidisciplinary care models and highlights the importance of considering both pharmacological and nutritional dimensions of patient care.

Establishing systematic vitamin monitoring and supplementation strategies may represent a promising approach in the comprehensive management of NP, although further evidence is needed to support their routine implementation. Ultimately, these findings highlight the interconnected nature of drug therapy, treatment duration, and nutritional status, and support the need for further research in this area. Due to the retrospective design, causality cannot be established, and the findings should be interpreted as associations.

Ethics

Ethics Committee Approval: Ethics committee approval was obtained on 8 January 2025 (University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, no: 85, date: 08.01.2025), and data extraction as well as statistical analysis were initiated only after this approval, in accordance with national regulations for retrospective data use.

Informed Consent: As this was a retrospective study based on previously recorded clinical data, the requirement for obtaining informed consent was waived by the ethics committee in accordance with institutional policy and national regulations. The study used anonymized data from routine clinical practice, with no direct patient contact or additional tests, and was conducted in accordance with national regulations governing retrospective chart reviews. No direct patient contact occurred, and no identifiable personal information was used.

Footnotes

Authorship Contributions

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

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

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

References

1. Zhang Y, Mustieles V, Wang YX, Sun Y, Agudelo J, Bibi Z, et al. Folate concentrations and serum perfluoroalkyl and polyfluoroalkyl substance concentrations in adolescents and adults in the USA (National Health and Nutrition Examination Study 2003-16): an observational study. Lancet Planet Health. 2023;7(6):e449-e458.
2. Bouhassira D, Attal N. Translational neuropathic pain research: a clinical perspective. Neuroscience. 2016;338:27-35.
3. Rosner J, Attal N, Finnerup NB. Clinical pharmacology of neuropathic pain. Int Rev Neurobiol. 2024;179:403-430.
4. Ma YC, Kang ZB, Shi YQ, Ji WY, Zhou WM, Nan W. The complexity of neuropathic pain and central sensitization: exploring mechanisms and therapeutic prospects. J Integr Neurosci. 2024;23(5):89.
5. Soliman N, Kersebaum D, Lawn T, Sachau J, Sendel M, Vollert J. Improving neuropathic pain treatment - by rigorous stratification from bench to bedside. J Neurochem. 2024;168(11):3699-3714.
6. Rahman MM, Jo YY, Kim YH, Park CK. Current insights and therapeutic strategies for targeting TRPV1 in neuropathic pain management. Life Sci. 2024;355:122954.
7. Testa L, Dotta S, Vercelli A, Marvaldi L. Communicating pain: emerging axonal signaling in peripheral neuropathic pain. Front Neuroanat. 2024;18:1398400.
8. Shkodina AD, Bardhan M, Chopra H, Anyagwa OE, Pinchuk VA, Hryn KV, et al. Pharmacological and non-pharmacological approaches for the management of neuropathic pain in multiple sclerosis. CNS Drugs. 2024;38(3):205-224.
9. Boccella S, De Filippis L, Giorgio C, Brandolini L, Jones M, Novelli R, et al. Combination drug therapy for the management of chronic neuropathic pain. Biomolecules. 2023;13(12):1802.
10. Andrejic N, Bozovic I, Moradi H, Tataei R, Knezevic NN. Neuropathic pain management: a focused review of current treatments and novel data from main ongoing clinical trials. Expert Opin Investig Drugs. 2025;34(4):287-299.
11. Tesfaye S, Sloan G, Petrie J, White D, Bradburn M, Young T, et al. Optimal pharmacotherapy pathway in adults with diabetic peripheral neuropathic pain: the OPTION-DM RCT. Health Technol Assess. 2022;26(39):1-100.
12. Liu X, Wei L, Zeng Q, Lin K, Zhang J. The treatment of topical drugs for postherpetic neuralgia: a network meta-analysis. Pain Physician. 2020;23(6):541-551.
13. Moisset X, Bouhassira D, Attal N. French guidelines for neuropathic pain: an update and commentary. Rev Neurol (Paris). 2021;177(7):834-837.
14. Dwajani S, Ravi L, Bindu CM, Prema R. Effect of antiepileptic drugs on serum vitamin B12 and folic acid among children with epilepsy in Bangalore. Biomedical and Pharmacology Journal. 2024;17(1):263-269.
15. Moisset X. Neuropathic pain: evidence based recommendations. Presse Med. 2024;53(2):104232.
16. Freynhagen R, Baron R, Huygen F, Perrot S. Narrative review of the efficacy and safety of the high-concentration (179mg) capsaicin patch in peripheral neuropathic pain with recommendations for clinical practice and future research. Pain Rep. 2025;10(2):e1235.

17. Kathiravan M, Kavitha S, Shanthi R. To determine the effect of long-term antiepileptic drug on the serum folate and vitamin B12 among epileptic patients. Sci Rep. 2021;11(1):4393.

18. Petroianu GA, Aloum L, Adem A. Neuropathic pain: mechanisms and therapeutic strategies. Front Cell Dev Biol. 2023;11:1072629.

19. Safahani M, Aligholi H, Asadi-Pooya AA. Management of antiepileptic drug-induced nutrition-related adverse effects. Neurol Sci. 2020;41(12):3491-3502.

20. Huang HL, Zhou H, Wang N, Yu CY. Effects of antiepileptic drugs on the serum folate and vitamin B12 in various epileptic patients. Biomed Rep. 2016;5(4):413-416.

21. Xu Y, Zhang N, Xu S, Xu H, Chen S, Xia Z. Effects of phenytoin on serum levels of homocysteine, vitamin B12, folate in patients with epilepsy: a systematic review and meta-analysis (PRISMA-compliant article). Medicine (Baltimore). 2019;98(12):e14844.

22. Becit-Kizilkaya M, Oncu S, Bugra Koca H, Cavusoglu D. Evaluation of anti-seizure medications and their serum concentration with regard to cardiovascular risk parameters: a cross-sectional study. Epilepsy Behav. 2024;155:109802.

23. O'Leary F, Samman S. Vitamin B12 in health and disease. Nutrients. 2010;2(3):299-316.

24. Hunt A, Harrington D, Robinson S. Vitamin B12 deficiency. BMJ. 2014;349:5226-5231.

25. Jones KS, Collins D, Meadows SR, Koulman A, Page P. National Diet and Nutrition Survey data reveal a decline in folate status in the United Kingdom population between 2008 and 2019. Am J Clin Nutr. 2023;118(6):1182-1191.

26. Holick MF, Binkley NC, Bischoff-Ferrari HA, Gordon CM, Hanley DA, Heaney RP, et al.; Endocrine Society. Evaluation, treatment, and prevention of vitamin D deficiency: an Endocrine Society clinical practice guideline. J Clin Endocrinol Metab. 2011;96(7):1911-1930. Erratum in: J Clin Endocrinol Metab. 2011;96(12):3908. Erratum in: J Clin Endocrinol Metab. 2024;109(10):e1991.

27. Seğmen F, Aydemir S, Küçük O, Dokuyucu R. The roles of vitamin D levels, gla-rich protein (GRP) and matrix gla protein (MGP), and inflammatory markers in predicting mortality in intensive care patients: a new biomarker link? Metabolites. 2024;14(11):620.

28. Julian T, Syeed R, Glascow N, Angelopoulou E, Zis P. B12 as a treatment for peripheral neuropathic pain: a systematic review. Nutrients. 2020;12(8):2221.

29. Linnebank M, Moskau S, Semmler A, Widman G, Stoffel-Wagner B, Weller M, et al. Antiepileptic drugs interact with folate and vitamin B12 serum levels. Ann Neurol. 2011;69(2):352-359.

30. Reyes Fernandez PC, Wright CS, Warden SJ, Hum J, Farach-Carson MC, Thompson WR. Effects of gabapentin and pregabalin on calcium homeostasis: implications for physical rehabilitation of musculoskeletal tissues. Curr Osteoporos Rep. 2022;20(6):365-378.

31. Alles SRA, Cain SM, Snutch TP. Pregabalin as a pain therapeutic: beyond calcium channels. Front Cell Neurosci. 2020;14:83.

32. He Q, Li J. The evolution of folate supplementation - from one size for all to personalized, precision, poly-paths. J Transl Int Med. 2023;11(2):128-137.

33. Li Y, Bromley R, Meador KJ, Bjork MH, Pennell PB, Voinescu PE; Epilepsy in the Childbearing Age through Menopause Consortium (ECAM). Folate supplement dosage for pregnant people with epilepsy: survey of international experts recommendation and current practices. Epileptic Disord. 2025;27(6):1209-1216.

34. Eghrari NB, Yazji IH, Yavari B, Van Acker GM, Kim CH. Risk of dementia following gabapentin prescription in chronic low back pain patients. Reg Anesth Pain Med. 2025:rapm-2025-106577. Epub ahead of print.

35. Mohan A, Kumar R, Kumar V, Yadav M. Homocysteine, vitamin B12 and folate level: possible risk factors in the progression of chronic heart and kidney disorders. Curr Cardiol Rev. 2023;19(4):e090223213539.

36. Sirivarasai J, Shantavasinkul PC, Thitiwiwatkul M, Monsuwan W, Panpunuan P, Sritara P. Association between plasma homocysteine, folate, vitamin B12 levels, and metabolic dysfunction indices in elderly with arterial stiffness. J Clin Med. 2025;14(9):2998.

37. Qu R, Han A, Zhao Y, Qu X, Zhu Y. Characteristics of clinical manifestations and molecular genetics of inherited hyperhomocysteinemia in children and adolescents: a single center experience from China. Front Neurol. 2025;16:1588273.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):290-296
DOI: 10.4274/BMB.galenos.2026.30075

Impact of COVID-19 Vaccination Status on Intensive Care Unit Mortality in Patients with COVID-19 Pneumonia: A Retrospective Observational Study

COVID-19 Pnömonisi Tanısıyla Yoğun Bakım Ünitesinde Yatan Hastalarda Aşılanma Durumunun Mortalite Üzerine Etkisi: Retrospektif Gözlemsel Çalışma

Sezen Kumaş Solak[1], Ceren Gür[2], Burcu Akyüz[1], Mehmet Salih Sevdi[1], Serdar Demirgan[1], Rasim Onur Karaoğlu[1], Ayşin Selcan[1]

[1]University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Anesthesiology and Reanimation, İstanbul, Turkey
[2]University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Internal Medicine, İstanbul, Turkey

Abstract

Objective: The aim of this study is to evaluate the effects of age, gender, clinical scores [sequential organ failure assessment (SOFA), acute physiology and chronic health evaluation (APACHE) II], radiological involvement, vaccination status, need for intubation, presence of comorbidities, waiting times at home and in the hospital ward, and the general clinical condition at the time of intensive care unit (ICU) admission on mortality in patients who were followed up in the intensive care unit due to coronavirus disease-2019 (COVID-19).

Method: Data from 806 patients with confirmed COVID-19 pneumonia admitted to the Intensive Care Unit of the University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital (March 2020-September 2021) were retrospectively analyzed. Demographic, clinical, and radiological data—including vaccination status (BioNTech or Sinovac), number of doses, and intubation status—were recorded. Independent predictors of mortality were identified using multivariable logistic regression analysis.

Öz

Amaç: Bu çalışmanın amacı, koronavirüs hastalığı-2019 (COVID-19) nedeniyle yoğun bakımda takip edilen hastalarda yaş, cinsiyet, klinik skorlar [ardışık organ yetmezliği değerlendirmesi (SOFA), akut fizyoloji ve kronik sağlık değerlendirmesi (APACHE) II], radyolojik tutulum, aşılanma durumu, entübasyon gereksinimi, komorbidite varlığı, evde ve serviste bekleme süreleri ile yoğun bakım ünitesi (YBÜ) kabul anındaki genel durumun mortalite üzerindeki etkilerini değerlendirmektir.

Yöntem: Mart 2020-Eylül 2021 tarihleri arasında Sağlık Bilimleri Üniversitesi, İstanbul Bağcılar Eğitim ve Araştırma Hastanesi Yoğun Bakım Ünitesi'nde COVID-19 pnömonisi tanısıyla izlenen 806 hastanın verileri retrospektif olarak incelendi. Hastaların demografik, klinik ve radyolojik bulguları; aşılanma durumu ve tipi (BioNTech, Sinovac), doz sayısı, entübasyon gereksinimi, SOFA ve APACHE II skorları kaydedildi. Mortaliteyi etkileyen faktörler çok değişkenli lojistik regresyon analizi ile değerlendirildi.

Address for Correspondence: Sezen Kumaş Solak, MD, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Anesthesiology and Reanimation, İstanbul, Turkey
E-mail: sezenkumassolak@gmail.com **ORCID:** orcid.org/0000-0002-9856-6269
Received: 06.11.2025 **Accepted:** 24.02.2026 **Epub:** 02.03.2026 **Publication Date:** 29.09.2026

Cite this article as: Kumaş Solak S, Gür C, Akyüz B, Sevdi MS, Demirgan S, Karaoğlu RO, et al. Impact of COVID-19 vaccination status on intensive care unit mortality in patients with COVID-19 pneumonia: a retrospective observational study. Bagcilar Med Bull. 2026;11(3):290-296

Abstract

Results: Of 806 patients, 58.6% were male, and the mean age was 63.9±15.3 years. The overall mortality rate was 67.5%. Mortality was significantly associated with older age, higher APACHE II scores, need for intubation, poor general condition, comorbidities, and severe computed tomography (CT) involvement ($p<0.05$); the difference in SOFA scores between survivors and non-survivors did not reach statistical significance on recalculation. Mortality was 70.7% among unvaccinated and 48.3% among vaccinated patients. Patients vaccinated with BioNTech (BNT162b2) had a numerically lower mortality rate than those who received CoronaVac (Sinovac) (42.1% vs. 51.3%); however, this difference did not reach statistical significance (χ^2 test, $p\approx0.256$). In multivariate analysis, age, APACHE II score, intubation, poor general condition, comorbidities, and severe CT findings were independent risk factors for mortality, whereas vaccination was a protective factor (odds ratio=0.46; 95% confidence interval: 0.29-0.73; p=0.001).

Conclusion: Mortality in ICU patients with COVID-19 was independently associated with advanced age, comorbidities, poor clinical condition, higher APACHE II scores, intubation, and severe radiological involvement. Vaccination—particularly with the mRNA-based BioNTech vaccination was associated with lower mortality, underscoring the importance of early diagnosis, timely critical care, and widespread immunization in reducing ICU deaths.

Keywords: BioNTech, comorbidity, COVID-19, intensive care, intubation, mortality, Sinovac, vaccination

Öz

Bulgular: Hastaların %58,6'sı erkek, yaş ortalaması 63,9±15,3 yıl idi. Genel mortalite oranı %67,5 bulundu. Mortalite ileri yaş, yüksek APACHE II skoru, entübasyon gereksinimi, kötü genel durum, komorbidite ve ağır bilgisayarlı tomografi (BT) tutulumuyla anlamlı şekilde ilişkiliydi ($p<0,05$). Aşısız hastalarda mortalite oranı %70,7 iken, aşılılarda %48,3 olarak saptandı. BioNTech (BNT162b2) aşısı yapılanlarda mortalite oranı, CoronaVac (Sinovac) grubuna göre belirgin olarak daha düşüktü (%41,2'ye karşı %55,6, p=0,038). Lojistik regresyon analizine göre yaş, APACHE II skoru, entübasyon gereksinimi, kötü genel durum, komorbidite varlığı ve ağır BT tutulumu mortaliteyi artıran bağımsız değişkenler; aşılanma ise koruyucu faktördü (odds oranı=0,46; %95 güven aralığı: 0,29-0,73; p=0,001).

Sonuç: COVID-19 hastalarında mortaliteyi belirleyen başlıca faktörler ileri yaş, komorbidite, kötü klinik durum, yüksek APACHE II skoru, entübasyon ve ağır radyolojik tutulumdur. Aşılanma, özellikle mRNA temelli BioNTech aşısı, YBÜ mortalitesini anlamlı düzeyde azaltmaktadır. Erken başvuru, hızlı klinik değerlendirme ve yaygın aşılanma daha düşük mortaliteyle ilişkili bulundu.

Anahtar kelimeler: Aşı, BioNTech, COVID-19, entübasyon, komorbidite, mortalite, Sinovac, yoğun bakım

Introduction

Since its emergence in Wuhan, China, in late 2019, coronavirus disease-2019 (COVID-19) has evolved into a major global public health crisis and has been declared a pandemic by the World Health Organization on March 11, 2020 (1). The disease presents with a wide clinical spectrum ranging from asymptomatic infection to severe pneumonia, acute respiratory distress syndrome, and multiorgan failure requiring intensive care support.

Advanced age, male sex, obesity, hypertension, diabetes mellitus, cardiovascular disease, and chronic kidney failure have been consistently identified as major risk factors for severe disease and mortality (2-4).

Despite advances in supportive care, mortality among intensive care unit (ICU) patients—particularly those requiring mechanical ventilation—remains high, ranging from 30% to 70% worldwide (4,5). Disease severity scores, comorbidity burden, radiological lung involvement, and timing of treatment initiation are key determinants of outcome (6-10).

Delayed access to healthcare, reflected by prolonged home and ward waiting times, has been associated with worsening hypoxemia, cytokine storm, and increased ICU mortality (11). Therefore, early diagnosis, timely referral, and prompt treatment initiation remain essential strategies to improve survival (10-12).

Vaccination has been the most effective intervention for reducing severe disease, hospitalization, and mortality (13-17). In Turkey, vaccination began with CoronaVac in January 2021, expanded to include BioNTech in April 2021, and booster recommendations were issued in July 2021. However, variations in vaccination coverage and hesitancy have influenced overall effectiveness.

Although international studies clearly show higher ICU admission and mortality rates among unvaccinated individuals (13-17), data focusing on critically ill COVID-19 patients in Turkey remain limited. Therefore, evaluating the impact of vaccination status on ICU mortality is essential to better define its protective role in this high-risk population.

Materials and Methods

This retrospective, observational cohort study was conducted in the ICU of the University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, between March 11, 2020, and September 30, 2021. Ethical approval was obtained from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Prof.

Dr. Cemil Taşcıoğlu City Hospital (approval no: E-48670771-514.99, date: 08.11.2021). Written informed consent was obtained from all patients (or their legal representatives, where applicable) prior to ICU admission and data collection. Patient data were retrospectively obtained from the hospital's electronic medical record system.

All patients admitted to the ICU during the study period with a confirmed diagnosis of COVID-19 pneumonia were included. Patient data were retrospectively obtained from the hospital's electronic medical record system.

COVID-19 pneumonia was diagnosed based on the presence of at least one respiratory symptom (fever, cough, or dyspnea) together with either a positive SARS-CoV-2 real-time reverse transcription polymerase chain reaction (RT-PCR) test or typical thoracic computed tomography (CT) findings.

Patients with negative RT-PCR results but thoracic CT findings compatible with COVID-19, such as bilateral peripheral ground-glass opacities, consolidations, and interstitial involvement, were also considered to have COVID-19 pneumonia. The diagnosis was confirmed through multidisciplinary clinical evaluation by infectious disease and intensive care specialists.

Demographic and clinical variables extracted from the electronic records included age, sex, comorbidities, degree of thoracic CT involvement, general clinical condition at ICU admission, time spent waiting at home, ward length of stay prior to ICU admission, ICU length of stay, need for endotracheal intubation, and clinical outcome (survival or death).

Thoracic CT involvement was graded by the attending radiologist according to the estimated percentage of lung parenchymal involvement as:

- Mild: ≤25%
- Moderate: 26-50%
- Severe: >50%.

General clinical condition at ICU admission was classified by the attending intensivist based on vital signs, oxygen requirement, level of consciousness, and hemodynamic stability into three categories:

- Good: Patients breathing spontaneously with low-flow oxygen support and hemodynamically stable.
- Moderate: Patients requiring high-flow oxygen therapy or non-invasive ventilation without vasopressor support.
- Poor: Patients requiring invasive mechanical ventilation and/or presenting with severe hypoxemia or hemodynamic instability.

Home waiting time was defined as the interval between the onset of the first COVID-19-related symptoms and the date of the first hospital admission, expressed in days. Symptom onset was recorded based on reports from the patient or family, as documented in the medical records. Ward waiting time was defined as the duration between hospital admission and ICU transfer.

Vaccination status and vaccination dates were verified using the Turkish Ministry of Health National Vaccination Tracking System. During the national vaccination program, two vaccines were available: the inactivated CoronaVac (Sinovac) and the mRNA-based BNT162b2 (BioNTech).

Patients were categorized into eight predefined groups according to vaccine type, number of doses, and immune maturation interval:

0 = Unvaccinated
1 = One dose of Sinovac
2 = Two doses of Sinovac
3 = Three doses of Sinovac
4 = Two doses of Sinovac + one dose of BioNTech
5 = One dose of BioNTech
6 = Two doses of BioNTech
7 = Vaccinated but <14 days since the last dose (immune response not yet developed).

For the purposes of statistical analysis, patients in category 7 (vaccinated but with less than 14 days elapsed since the last dose) were classified together with unvaccinated patients, as protective immunity was not yet considered established at the time of ICU admission.

A total of 997 patients were screened for eligibility. Of these, 191 were excluded because of missing data regarding vaccination status, thoracic CT findings, general clinical condition, or clinical outcome. The final analysis included 806 patients.

As the study encompassed all eligible ICU admissions during the specified period, no prior sample size calculation was performed. Continuous variables were checked for plausibility, and clearly erroneous values were treated as missing data. Only complete cases were included in regression analyses.

Statistical Analysis

All analyses were conducted using IBM SPSS Statistics version 25.0 (IBM Corp., Armonk, NY, USA). Continuous

variables were tested for normality using the Shapiro-Wilk test. Normally distributed variables were expressed as mean ± standard deviation, and non-normally distributed variables as median [interquartile range, (IQR)]. Categorical variables were summarized as numbers (percentages). Comparisons between survivors and non-survivors were made using the Student's t-test or Mann-Whitney U test for continuous data and the chi-square or Fisher's exact test for categorical data. To identify independent predictors of mortality, univariate logistic regression analyses were first performed; variables with $p<0.10$ were subsequently entered into a multivariable logistic regression model using backward stepwise selection. Adjusted odds ratios and 95% confidence intervals were reported. A two-tailed $p<0.05$ was considered statistically significant.

Results

Demographic and Baseline Characteristics

A total of 806 patients with COVID-19 pneumonia were included in the final analysis. The cohort consisted of 472 males (58.6%) and 334 females (41.4%), with a mean age of 63.9±15.3 years (range: 16-102). The overall ICU mortality rate was 67.5% (n=544). Baseline demographic and clinical characteristics of the study population are summarized in Table 1.

The mean age of non-survivors was significantly higher than that of survivors (66.7±14.9 vs. 58.2±15.8 years, $p<0.001$). At ICU admission, 596 patients (73.9%) were classified as having poor general condition, 195 (24.2%) as moderate, and 15 (1.9%) as good. Mortality rates were 486/596 (81.5%) for poor, 56/195 (28.7%) for moderate, and 2/15 (13.3%) for good general condition; the sum of deaths across these three groups (544) corresponds exactly to overall ICU mortality ($p<0.001$).

Clinical and Radiological Findings

Comorbidities were present in 85.7% of non-survivors compared with 61.2% of survivors ($p<0.001$). Comparisons of major clinical variables between survivors and non-survivors are shown in Table 2.

Mean sequential organ failure assessment (SOFA) score was similar between non-survivors and survivors (13.7±3.8 vs. 13.4±3.2); this difference was not statistically significant on recalculation ($p\approx0.270$). While the acute physiology and chronic health evaluation (APACHE) II score was significantly higher (27.8±6.9 vs. 25.1±5.7, $p<0.001$). The median ICU length of stay was also longer among non-survivors [11.2 (IQR 6-18) vs. 9.4 (IQR 5-16) days, $p=0.032$].

Thoracic CT involvement was classified as mild in 132 patients (16.4%), moderate in 272 (33.8%), and severe in 402 (49.8%). Mortality increased significantly with increasing CT severity ($p<0.001$) (Table 3).

Vaccination Status and Outcomes

Among all patients, 632 (78.4%) were unvaccinated, whereas 174 (21.6%) had received at least one dose of a COVID-19 vaccine. Vaccination status and associated mortality rates are presented in Table 4.

The mortality rate was significantly higher in unvaccinated patients than in vaccinated patients (70.7% vs. 48.3%, $p<0.001$). Among vaccinated patients, those who received BioNTech exhibited numerically lower mortality than those vaccinated with CoronaVac (42.1% vs. 51.3%); however, this difference was not statistically significant ($p\approx0.256$).

Predictors of Mortality

Multivariable logistic regression analysis identified several independent predictors of ICU mortality (Table 5).

Vaccination was identified as an independent protective factor against ICU mortality.

Table 1. Demographic and baseline clinical characteristics of ICU patients with COVID-19

Variable	Total (n=806)
Age, mean ± SD (years)	63.9±15.3
Male, n (%)	472 (58.6)
Female, n (%)	334 (41.4)
Poor general condition, n (%)	596 (73.9)
Moderate general condition, n (%)	195 (24.2)
Good general condition, n (%)	15 (1.9)
Intubation requirement, n (%)	559 (69.4)
ICU mortality, n (%)	544 (67.5)

SD: Standard deviation, ICU: Intensive care unit, COVID-19: Coronavirus disease-2019

Table 2. Clinical characteristics according to survival status

Variable	Survivors	Non-survivors	p-value
Age (years), mean ± SD	58.2±15.8	66.7±14.9	<0.001
Comorbidity ≥1, %	61.2	85.7	<0.001
SOFA score, mean ± SD	13.4±3.2	13.7±3.8	0.047
APACHE II score, mean ± SD	25.1±5.7	27.8±6.9	<0.001
ICU-LOS, median (IQR)	9.4 (5-16)	11.2 (6-18)	0.032

SD: Standard deviation, ICU: Intensive care unit, IQR: Interquartile range, SOFA: Sequential organ failure assessment, LOS: Length of stay, APACHE: Acute physiology and chronic health evaluation

Table 3. Thoracic CT involvement and mortality

CT involvement	Total, n (%)	Mortality, n (%)
Mild (≤25%)*	132 (16.4)	64 (48.5)
Moderate (26-50%)	272 (33.8)	149 (54.8)
Severe (>50%)	402 (49.9)	331 (82.4)
Total	806 (100.0)	544 (67.5)

*: The mild involvement group includes 23 patients with a confirmed COVID-19 pneumonia diagnosis by clinical/RT-PCR criteria but without detectable CT involvement.
COVID-19: Coronavirus disease-2019, RT-PCR: Real time-polymerase chain reaction, CT: Computed tomography

Table 4. Vaccination status and mortality

Vaccination status	n (%)	Mortality, n (%)
Unvaccinated	632 (78.4)	447 (70.7)
Vaccinated (any dose)	174 (21.6)	84 (48.3)
CoronaVac only	117 (14.5)	65 (55.6)
BioNTech (any dose)	57 (7.1)	24 (41.2)

Table 5. Multivariable logistic regression analysis for ICU mortality

Variable	Adjusted OR	95% CI	p-value
Age (per year)	1.04	1.02-1.06	<0.001
Poor general condition	3.89	2.12-7.13	<0.001
Severe CT involvement	2.77	1.46-5.26	0.002
≥1 comorbidity	2.23	1.41-3.56	<0.001
Intubation requirement	4.62	2.87-7.44	<0.001
APACHE II score	1.08	1.04-1.12	<0.001
Vaccinated (any dose)	0.46	0.29-0.73	0.001

OR: Odds ratio, CI: Confidence interval, ICU: Intensive care unit, APACHE: Acute physiology and chronic health evaluation, CT: Computed tomography

Summary of Key Findings

Mortality was significantly associated with advanced age, poor general clinical condition, comorbidities, high APACHE II scores, severe thoracic CT involvement, and need for intubation. Vaccination was an independent protective factor that significantly reduced mortality risk.

Discussion

In this retrospective cohort of 806 critically ill patients with COVID-19, we comprehensively evaluated the factors influencing ICU mortality. Our findings demonstrated that advanced age, the presence of comorbidities, poor general clinical condition at ICU admission, higher APACHE II scores, the requirement for endotracheal intubation, and severe thoracic involvement on CT were independently associated with increased mortality. Conversely, vaccination (with either vaccine) was significantly associated with reduced risk of death; the difference between vaccine types (BioNTech vs. CoronaVac) did not reach statistical significance in this cohort. These results are consistent with previous multicenter ICU studies conducted worldwide (5).

Age, Comorbidities, and Disease Severity

Since the onset of the pandemic, older age and comorbidities have consistently been identified as major risk factors for severe COVID-19 and mortality (6-8). In the landmark study by Zhou et al. (4), age, hypertension, and diabetes were found to be independent predictors of poor outcomes, a finding later supported by Petrilli et al. (6) and Jordan et al. (7). In our cohort, non-survivors were significantly older and exhibited a markedly higher prevalence of chronic diseases (85.7%).

Each one-year increase in age was associated with a 4% increase in the odds of death. Underscoring the impact of immunosenescence and reduced physiological reserve in elderly patients. Poor general condition at ICU admission was associated with an 81.5% mortality rate, suggesting that delayed presentation and pre-ICU clinical deterioration play critical roles in disease progression.

Both APACHE II and SOFA scores have been validated as objective predictors of disease severity in critically ill COVID-19 patients. In the present study, higher APACHE II scores remained independently associated with mortality, supporting previous reports linking multiorgan dysfunction and systemic inflammation to fatal outcomes (4,12). Similarly, patients requiring mechanical ventilation had a more than fourfold increased mortality odds, consistent with other ICU cohorts (10).

Prognostic Value of Thoracic CT Involvement

Radiological assessment of pulmonary involvement is an essential component of COVID-19 severity evaluation. We observed a strong stepwise relationship between CT involvement and mortality, with 31% in mild disease, 54.8% in moderate disease, and 82.4% in severe disease. These findings align with those of Wu et al. (11) and Francone et al. (9), who demonstrated that higher CT scores correlated with worsening oxygenation and increased ventilatory support requirements. Extensive radiological involvement likely reflects diffuse alveolar damage and a heightened inflammatory burden, leading to severe hypoxemia and a poor prognosis. Accordingly, CT severity scoring may serve as a valuable complementary prognostic tool alongside clinical severity indices in early risk stratification.

Impact of Vaccination on Mortality

The protective effect of vaccination against severe COVID-19 and death has been clearly demonstrated

worldwide. In the present study, mortality was 70.7% among unvaccinated patients compared with 48.3% among vaccinated individuals ($p<0.001$). Among vaccinated patients, mortality was numerically lower in those who received BioNTech (mRNA) than in those who received CoronaVac (42.1% vs. 51.3%), although this difference did not reach statistical significance ($p\approx0.256$).

This observation is consistent with the findings of Haas et al. (13), who reported a marked reduction in COVID-19-related mortality following widespread BioNTech vaccination in Israel. Tanriover et al. (14) demonstrated that CoronaVac reduced severe disease but elicited a shorter-lived immune response. These findings are consistent with a randomized controlled phase 3 trial demonstrating robust immunogenicity of mRNA vaccines [Polack et al. (15)] and with large-scale observational studies demonstrating reduced risk of hospitalization under real-world conditions [Sheikh et al. (16); Tartof et al. (17)]. As none of these three studies directly evaluated booster-dose effectiveness, the statement regarding improved survival with booster doses is not supported by these citations and should either be removed or replaced with an appropriate reference specifically addressing booster-dose effectiveness.

Our results reinforce the importance of sustained vaccination strategies and booster administration to mitigate ICU mortality in high-risk groups.

Timing of Care and Delayed Admission

As observed by Li et al. (12), delayed presentation permits progression of hypoxemia and cytokine storm, thereby reducing the opportunity for effective intervention. Early diagnosis, timely hospital admission, and prompt initiation of respiratory support are widely regarded as crucial determinants of survival in severe COVID-19.

Study Limitations

The strengths of this study include its large sample size, reliable vaccination data verified through the National Vaccination Tracking System, and standardized classification of thoracic CT involvement. However, several limitations should be acknowledged.

The retrospective design and heterogeneity of treatment protocols may introduce residual confounding. In addition, viral variant differences and the absence of certain inflammatory biomarkers (such as IL-6, D-dimer, and ferritin) limit external generalizability. Despite these limitations, the study provides robust real-world evidence regarding mortality determinants in vaccinated and unvaccinated ICU patients with COVID-19.

Clinical Implications

Our findings emphasize the multifactorial nature of COVID-19 mortality and the necessity for integrated management strategies. Patients presenting with advanced age, comorbidities, high APACHE II scores, or severe CT involvement require early intensive monitoring and aggressive supportive care. Moreover, maintaining high vaccination and booster coverage remains a critical, cost-effective approach to reducing ICU admissions and mortality.

Conclusion

In conclusion, advanced age, comorbidities, poor general condition at ICU admission, high APACHE II scores, requirement for intubation, and severe thoracic CT involvement were identified as strong independent predictors of ICU mortality in patients with COVID-19.

Conversely, vaccination was significantly associated with reduced mortality, underscoring its protective role in high-risk individuals, and underscores the importance of early recognition, timely referral, and rapid initiation of critical care.

These findings demonstrate that a comprehensive strategy combining early intervention, optimized intensive care management, and sustained vaccination coverage is essential for improving survival in patients with severe COVID-19.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Prof. Dr. Cemil Taşcıoğlu City Hospital (approval no: E-48670771-514.99, date: 08.11.2021).

Informed Consent: Written informed consent was obtained from all patients (or their legal representatives, where applicable) prior to ICU admission and data collection. Patient data were retrospectively obtained from the hospital's electronic medical record system.

Footnotes

Authorship Contributions

Surgical and Medical Practices: S.K.S., C.G., B.A., M.S.S., Concept: S.K.S., S.D., R.O.K., A.S., Design: S.K.S., S.D., R.O.K.,

A.S., Data Collection or Processing: S.K.S., C.G., B.A., A.S., Analysis or Interpretation: S.K.S., C.G., B.A., S.D., R.O.K., Literature Search: S.K.S., C.G., B.A., S.D., R.O.K., Writing: S.K.S., B.A., S.D., C.G., R.O.K.

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

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

References

1. World Health Organization. WHO Director-General's opening remarks at the media briefing on COVID-19 – 11 March 2020. Geneva: World Health Organization; 2020.
2. Wu Z, McGoogan JM. Characteristics of and important lessons from the coronavirus disease 2019 outbreak in China. JAMA. 2020;323(13):1239-1242.
3. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med. 2020;382(18):1708-1720.
4. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet. 2020;395(10229):1054-1062. Erratum in: Lancet. 2020;395(10229):1038.
5. Grasselli G, Zangrillo A, Zanella A, Antonelli M, Cabrini L, Castelli A, et al. Baseline characteristics and outcomes of 1591 patients infected with SARS-CoV-2 admitted to ICUs of the Lombardy region, Italy. JAMA. 2020;323(16):1574-1581. Erratum in: JAMA. 2021;325(20):2120.
6. Petrilli CM, Jones SA, Yang J, Rajagopalan H, O'Donnell L, Chernyak Y, et al. Factors associated with hospital admission and critical illness among 5279 people with coronavirus disease 2019 in New York City: prospective cohort study. BMJ. 2020;369:m1966.
7. Jordan RE, Adab P, Cheng KK. Covid-19: risk factors for severe disease and death. BMJ. 2020;368:m1198.
8. Liu X, Zhou H, Zhou Y, Wu X, Zhao Y, Lu Y, et al. Risk factors associated with disease severity and length of hospital stay in COVID-19 patients. J Infect. 2020;81(1):e95-e97.
9. Francone M, Iafrate F, Masci GM, Coco S, Cilia F, Manganaro L, et al. Chest CT score in COVID-19 patients: correlation with disease severity and short-term prognosis. Eur Radiol. 2020;30(12):6808-6817.
10. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet. 2020;395(10223):497-506. Erratum in: Lancet. 2020;395(10223):496.
11. Wu C, Chen X, Cai Y, Xia J, Zhou X, Xu S, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China. JAMA Intern Med. 2020;180(7):934-943. Erratum in: JAMA Intern Med. 2020;180(7):1031.
12. Li K, Wu J, Wu F, Guo D, Chen L, Fang Z, et al. The clinical and chest CT features associated with severe and critical COVID-19 pneumonia. Invest Radiol. 2020;55(6):327-331.
13. Haas EJ, Angulo FJ, McLaughlin JM, Anis E, Singer SR, Khan F, et al. Impact and effectiveness of mRNA BNT162b2 vaccine against SARS-CoV-2 infections and COVID-19 cases, hospitalisations, and deaths following a nationwide vaccination campaign in Israel: an observational study using national surveillance data. Lancet. 2021;397(10287):1819-1829. Erratum in: Lancet. 2021;398(10296):212.
14. Tanriover MD, Doğanay HL, Akova M, Güner HR, Azap A, Akhan S, et al. Efficacy and safety of an inactivated whole-virion SARS-CoV-2 vaccine (CoronaVac): interim results of a double-blind, randomised, placebo-controlled, phase 3 trial in Turkey. Lancet. 2021;398(10296):213-222. Erratum in: Lancet. 2022;399(10323):436.
15. Polack FP, Thomas SJ, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Safety and efficacy of the BNT162b2 mRNA Covid-19 vaccine. N Engl J Med. 2020;383(27):2603-2615.
16. Sheikh A, McMenamin J, Taylor B, Robertson C; Public Health Scotland and the EAVE II Collaborators. SARS-CoV-2 Delta VOC in Scotland: demographics, risk of hospital admission, and vaccine effectiveness. Lancet. 2021;397(10293):2461-2462.
17. Tartof SY, Slezak JM, Fischer H, Hong V, Ackerson BK, Ranasinghe ON, et al. Effectiveness of mRNA BNT162b2 COVID-19 vaccine up to 6 months in a large integrated health system in the USA: a retrospective cohort study. Lancet. 2021;398(10309):1407-1416.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):297-306
DOI: 10.4274/BMB.galenos.2026.06978

Maternal Immune Dysregulation in Preeclampsia: A Comparative Analysis of Maternal Serum Interferon Alpha-1 (IFN α-1) Concentrations in Preeclamptic and Healthy Pregnancies

Preeklampside Maternal İmmün Disregülasyon: İnterferon Alfa-1 (IFN α-1) Maternal Serum Konsantrasyonlarının Preeklamptik ve Sağlıklı Gebeliklerde Karşılaştırmalı Analizi

Barış Boza[1], Verda Alpay[1], Fırat Ersan[1], Hümeyra Öztürk Emre[2]

[1]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Obstetrics and Gynecology, Division of Perinatology, İstanbul, Turkey
[2]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Biochemistry, İstanbul, Turkey

Abstract

Objective: Preeclampsia is a serious pregnancy-related hypertensive disorder that is linked to disturbances in the immune system. Type I interferons (IFNs), especially IFN-alpha 1 (α-1), are thought to play a role in this condition, but their exact behavior in the general preeclamptic population is still unclear. In this study, we aimed to measure maternal serum IFN α-1 levels in women with early-onset preeclampsia and to assess whether IFN α-1 could be useful as a diagnostic biomarker.

Method: In this prospective case-control study, we included 100 pregnant women: 50 with early-onset preeclampsia and 50 healthy, normotensive controls matched for maternal age, gestational age, and body mass index. Maternal serum IFN α-1 levels were measured using an enzyme-linked immunosorbent assay. Group comparisons were conducted statistically, and a receiver operating characteristic (ROC) curve analysis was performed to evaluate the diagnostic accuracy of IFN α-1.

Results: Maternal serum IFN α-1 levels were significantly higher in the preeclampsia group than in the healthy controls (80.5±31.60 pmol/L vs. 57.9±23.32 pmol/L; p=0.001). However, despite this significant

Öz

Amaç: Preeklampsi, immün sistem bozukluklarıyla ilişkili ciddi bir gebeliğe özgü hipertansif hastalıktır. Tip I interferonların (IFNs), özellikle IFN-alfa 1'in (α-1), preeklampsinin patofizyolojisinde rol oynadığı düşünülmektedir. Bu çalışmada, erken başlangıçlı preeklampsisi olan kadınlarda maternal serum IFN α-1 düzeylerinin değerlendirilmesi ve tanısal biyobelirteç olarak kullanılabilirliğinin araştırılması amaçlanmıştır.

Yöntem: Prospektif olgu-kontrol tasarımındaki bu çalışmaya, erken başlangıçlı preeklampsisi olan 50 gebe ve maternal yaş, gebelik haftası ve vücut kitle indeksi açısından eşleştirilmiş 50 sağlıklı, normotansif kontrol dahil edilmiştir. Maternal serum IFN α-1 düzeyleri enzim bağlantılı immünosorbent testi yöntemiyle ölçülmüş; gruplar istatistiksel olarak karşılaştırılmış ve tanısal doğruluğu değerlendirmek için alıcı çalışma karakteristiği (ROC) eğrisi analizi yapılmıştır.

Bulgular: Maternal serum IFN α-1 düzeyleri preeklampsi grubunda kontrol grubuna kıyasla anlamlı derecede yüksek bulunmuştur (80,5±31,60 pmol/L vs. 57,9±23,32 pmol/L; p=0,001). Ancak bu belirgin artışa rağmen, ROC eğrisi analizi sınırlı bir tanısal performans

Address for Correspondence: Barış Boza, MD, University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Obstetrics and Gynecology, Division of Perinatology, İstanbul, Turkey
E-mail: barisboza@hotmail.com **ORCID:** orcid.org/0000-0002-3831-3745
Received: 13.12.2025 **Accepted:** 02.05.2026 **Epub:** 06.05.2026 **Publication Date:** 29.09.2026

Cite this article as: Boza B, Alpay V, Ersan F, Öztürk Emre H. Maternal immune dysregulation in preeclampsia: a comparative analysis of maternal serum interferon alpha-1 (IFN α-1) concentrations in preeclamptic and healthy pregnancies. Bagcilar Med Bull. 2026;11(3):297-306

Abstract

elevation, the ROC curve analysis demonstrated limited diagnostic performance, with an area under the curve of 0.589 (95% confidence interval: 0.382-0.795). Using an optimal cut-off value of 80.15 pmol/L, IFN α-1 achieved a sensitivity of 61% and a specificity of 64%.

Conclusion: Maternal serum IFN α-1 levels are significantly elevated in women with early-onset preeclampsia, supporting the hypothesis that activation of the type I IFN pathway is an important contributor to the disease's pathophysiology. However, given its poor discriminatory performance, serum IFN α-1 appears to have limited value as a standalone diagnostic biomarker.

Keywords: Biomarker, IFN α-1, preeclampsia, type I interferons

Öz

göstermiş olup eğri altındaki alan 0,589 (%95 güven aralığı: 0,382-0,795) olarak bulunmuştur. 80,15 pmol/L optimal kesim değeri kullanıldığında, IFN α-1 için duyarlılık %61, özgüllük ise %64 olarak saptanmıştır.

Sonuç: Erken başlangıçlı preeklampside maternal serum IFN α-1 düzeyleri anlamlı olarak artmıştır ve bu bulgu tip I IFN yolaklarının hastalığın patofizyolojisindeki rolünü desteklemektedir. Bununla birlikte, IFN α-1'in tek başına tanısal biyobelirteç olarak kullanımı sınırlı olup, daha çok altta yatan enflamatuvar durumu yansıtan bir belirteç olabileceği düşünülmektedir.

Anahtar kelimeler: Biyobelirteç, IFN α-1, preeklampsi, tip I interferonlar

Introduction

Preeclampsia is a serious hypertensive disorder of pregnancy, affecting approximately 2-8% of pregnancies worldwide and remaining a major cause of maternal and perinatal morbidity and mortality (1,2). It is characterized by new-onset hypertension after 20 weeks of gestation, accompanied by proteinuria and/or signs of maternal organ dysfunction. Beyond its acute clinical impact, preeclampsia is associated with increased long-term cardiovascular and metabolic risks for both mothers and offspring. Despite advances in obstetric care, delivery of the placenta and fetus remains the only definitive treatment, often resulting in iatrogenic preterm birth (2).

Although the pathogenesis of preeclampsia is multifactorial and incompletely understood, placental dysfunction is widely regarded as a central mechanism. Impaired trophoblast invasion and inadequate remodeling of the maternal spiral arteries lead to reduced placental perfusion, hypoxia, and oxidative stress, triggering the release of placental-derived factors into the maternal circulation and contributing to the systemic manifestations of the disease (1,2).

A key feature of this process is an imbalance between anti-angiogenic and pro-angiogenic factors—specifically, increased levels of fms-like tyrosine kinase-1 (sFlt-1) and soluble endoglin, and a deficit of pro-angiogenic factors like placental growth factor (PlGF). This dysregulation contributes to widespread maternal endothelial dysfunction, promoting systemic inflammation, vasoconstriction, and ultimately multiorgan involvement. Together, these mechanisms give rise to the clinical manifestations of preeclampsia (2).

Growing evidence indicates that immunological maladaptation at the maternal-fetal interface is an important upstream mechanism contributing to abnormal placentation. A successful pregnancy requires the development of maternal immune tolerance toward the semi-allogeneic fetus, which is partially achieved through a shift from a pro-inflammatory T-helper 1 (Th1) response to a more anti-inflammatory T-helper 2 profile (2,3). In preeclampsia, this immunological balance is disrupted, with a reversion toward a Th1-dominant state marked by excessive production of pro-inflammatory cytokines such as tumor necrosis factor-alpha (TNF-α), interleukin-6 and interferons (IFN) (2,4). This heightened inflammatory milieu is thought to impair trophoblast invasion and promote endothelial dysfunction, both of which are central features of the disorder.

Several studies investigating interferon alpha-1 (IFN α-1) and vascular endothelial injury in systemic lupus erythematosus (SLE) have also shown that patients with elevated IFN α-1 levels exhibit greater endothelial damage. Higher IFN α-1 concentrations have been found to be associated with increased vascular endothelial injury and enhanced endothelial apoptosis in individuals with SLE (5,6).

During normal placentation, trophoblast cells invade the maternal uterine vasculature and remodel it by replacing the endothelial lining of the spiral arteries (7,8). Inadequate spiral artery transformation is a well-recognized hallmark of preeclampsia (9). In addition, *in vitro* studies on spiral artery transformation have shown that IFN α-1 and sFlt-1 levels interact with extravillous trophoblast cells (10,11).

In this context, the role of type I interferons (IFNs), especially IFN-α, has received growing attention. Type

I IFNs are a group of cytokines that are important for the body's first immune response to viral infections. They are also involved in the development of autoimmune diseases including SLE (12,13). It is well known that pregnant women with SLE have a higher risk of developing preeclampsia (12). Several studies have reported that increased IFN-α levels or a stronger type I *IFN* gene signature are linked to the development of preeclampsia in this high-risk group, sometimes even before the symptoms of the disease appear (13-15).

The mechanisms through which type I IFNs may contribute to the development of preeclampsia appear to be complex and involve several pathways. Both *in vitro* and *in vivo* studies have shown that IFN-α has strong anti-angiogenic effects, reduces the invasive ability of extravillous trophoblasts, and triggers a gene expression pattern similar to that seen in preeclampsia (13,16). In addition, IFN-α can make the maternal endothelium more sensitive to anti-angiogenic factors such as sFlt-1, which further worsens endothelial dysfunction and supports the systemic features of preeclampsia (12,13). These findings suggest that an overactive type I IFN system may represent an important link between immune imbalance and the clinical presentation of the disease.

While the association between type I IFNs and preeclampsia is increasingly recognized in the context of autoimmune disorders, the role of IFN-α in the broader population of women with preeclampsia remains less defined. It is unclear whether the elevation of IFN α-1 is a common feature of preeclampsia or is restricted to cases with underlying autoimmune or inflammatory conditions. Moreover, the potential utility of maternal serum IFN α-1 as a biomarker for the diagnosis or risk stratification of preeclampsia has not been thoroughly investigated. Therefore, this study was designed to compare the maternal serum levels of IFN α-1 in healthy pregnant women and those with early-onset preeclampsia, and to evaluate its potential as a serum biomarker for the diagnosis or management of this devastating pregnancy complication.

Materials and Methods

This prospective cross-sectional case-control study was conducted at the Perinatology Department of University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital. The study protocol was approved by the local Ethics Committee (approval no. 2024-330, date: 10.09.2025), and all procedures were performed in accordance with the Declaration of Helsinki. Written informed consent was obtained from all participants prior to study enrollment.

A total of 100 pregnant women were prospectively enrolled and divided into two groups: a case group of 50 patients diagnosed with early-onset preeclampsia, and a control group of 50 healthy, normotensive pregnant women. The control group was matched to the preeclampsia group in terms of gestational age, maternal age and body mass index (BMI).

Early-onset preeclampsia was diagnosed based on new-onset hypertension (systolic blood pressure ≥140 mmHg and/or diastolic blood pressure ≥90 mmHg on at least two measurements) occurring before 34 weeks of gestation, together with proteinuria (≥300 mg in a 24-hour urine sample or a spot urine protein-to-creatinine ratio ≥0.3 mg/mg).

The inclusion criteria for all participants were a maternal age between 18 and 45 years and a singleton pregnancy. Women were excluded if they had a multiple pregnancy, known fetal genetic or structural abnormalities, pre-existing autoimmune diseases, thyroid disorders, pre-gestational or gestational diabetes mellitus, chronic hypertension, chronic kidney or liver disease, malignancy, recent major surgery, active maternal or perinatal infections, obstetric cholestasis, or suspected chorioamnionitis.

Comprehensive demographic and clinical data were collected from all participants at the time of enrollment. These data included maternal age, body mass index (BMI), gravidity, parity, obstetric history, smoking status, and the presence of any chronic medical conditions. Fetal biometry measurements and Doppler ultrasonography findings, including umbilical artery and uterine artery pulsatility indices, were also recorded.

Venous blood samples were obtained once from each participant, either during a routine outpatient visit or upon hospital admission before delivery. Blood was drawn from the antecubital vein into gel-containing serum separator tubes. The samples were immediately taken to the central laboratory and allowed to clot at room temperature. Serum was separated by centrifugation at 3000 × g for 10 minutes. The serum aliquots were then stored at -80 °C and were exposed to no more than one freeze-thaw cycle prior to analysis.

Maternal serum levels of IFN α-1 were measured using a commercial enzyme-linked immunosorbent assay kit (catalog no. E0129Hu, BT-LAB, Shanghai, China) according to the manufacturer's instructions. All patient

and standard samples were analyzed in duplicate to improve measurement accuracy. Optical density values were read at 450 nm using a microplate reader, and sample concentrations were calculated from a standard curve generated with a four-parameter logistic regression model.

Statistical Analysis

All statistical analyses were conducted using IBM SPSS version 24.0 (Chicago, USA). To determine the appropriate statistical tests, the distribution of continuous variables was assessed by visual inspection of histograms, evaluation of Skewness and Kurtosis values, and application of the Kolmogorov-Smirnov and Shapiro-Wilk tests. Continuous variables that were normally distributed were presented as mean ± standard deviation and were compared using Student's t-test. For variables that were not normally distributed, comparisons were made using the Mann-Whitney U test, and the results were expressed as median (minimum-maximum). Categorical variables were expressed as percentage (%) and were compared using chi-square test and Fisher's exact test if expected cell count was <5. For the examination of correlations between the maternal IFN α-1 and other parameters, Pearson's and Spearman's correlation analyses were performed. A p-value of <0.05 was considered statistically significant. Receiver operating characteristic (ROC) analysis was performed to evaluate the discriminative ability of maternal IFN α-1 levels in distinguishing patients with early-onset preeclampsia from healthy controls and to determine the optimal cut-off value.

A formal a priori sample size calculation could not be performed due to the absence of previous studies reporting maternal serum IFN α-1 levels in early-onset preeclampsia within a non-autoimmune population. Therefore, a post-hoc power analysis was conducted to evaluate whether the achieved sample size was sufficient to detect the observed differences between groups.

Based on the observed data, a large effect size (Cohen's d=0.80) was obtained. Assuming this effect size, a two-sample Student's t-test with an alpha level of 0.05 and a statistical power of 0.95 indicated that a total sample size of 84 participants (42 per group) would be required. Accordingly, the sample size of the present study (n=100) was considered adequate.

Results

A total of 100 pregnant women were included in the study, comprising 50 women diagnosed with early-onset preeclampsia and 50 women with uncomplicated pregnancies. Maternal demographic characteristics and birth-related parameters are summarized in Table 1. The mean maternal age in the preeclampsia and control groups was 29.6±5.22 and 28.8±5.11 years, respectively, while the mean gestational age at blood sampling was 29.6±0.87 and 29.3±0.69 weeks, respectively. No statistical difference was observed between the preeclampsia and control groups in terms of both parameters (p=0.441, p=0.079). In addition, the mean BMI of the patients was 31.2 (±5.06) and 27.4

Table 1. The demographic characteristics and the perinatal outcomes of the study groups

		Preeclampsia group (n=50)	Control group (n=50)	Total (n=100)	p-values[a]
Age (years)[c]		29.6 (±5.22)	28.8 (±5.11)	29.2 (±5.15)	0.441
Gravida (n)[b]		1 (1-6)	2 (1-7)	2 (1-7)	**0.022**
Parity (n)[b]		1 (1-5)	1 (0-3)	1 (0-5)	**0.014**
BMI (kg/m²)[c]		31.2 (±5.06)	27.4 (±4.41)	29.2 (±5.01)	**0.001**
Mean arterial pressure[c]		104.4 (±9.14)	80.5 (±6.26)	92.5 (±14.34)	0.001
GA at delivery (weeks)[c]		34.3 (±2.07)	38.4 (±1.53)	36.4 (±2.75)	**0.001**
Birth weight (g)[c]		1950 (±616)	3313 (±423)	2631 (±863)	**0.001**
1-minute Apgar score[b]		7 (4-8)	8 (6-9)	8 (4-9)	**0.001**
5-minute Apgar score[b]		8 (3-9)	9 (8-10)	9 (3-10)	**0.001**
Birth method[d]	**SVB**	3 (6%)	20 (40%)	23 (23%)	**0.001**
	CS	47 (94%)	30 (60%)	77 (77%)	
Gender[d]	**Female**	25 (50%)	28 (56%)	53 (53%)	0.548
	Male	25 (50%)	22 (44%)	47 (47%)	

BMI: Body mass index, GA: Gestational age, CS: Caesarean section, SVB: Spontaneous vaginal birth, [a]: Level of significance p<0.05, [b]: Data that were not normally distributed are expressed as median (minimum-maximum) and were compared using Mann-Whitney U test, [c]: Continuous variables distributed normally are expressed as mean and standard deviation and are compared using Student's t-test, [d]: Categorical variables were presented as percentage (%) and count (n) and were compared using the chi-square and Fisher's exact test

(±4.41) in the study and control groups, respectively, with a statistically significant difference between the study groups (p=0.001). In line with the literature, obstetric histories of the participants including gravidity and parity were lower in the preeclampsia group compared with women with normal pregnancies (p=0.022, p=0.014). Mean arterial pressures of the patients were 104.4 (±9.14), 80.5 (±6.26) in preeclampsia and control groups respectively. And a statistically significant difference was observed in terms of this parameter (p<0.001). Mean gestational age at delivery, and birth weight were 34.3 (±2.07), 38.4 (±1.53) weeks, 1950 (±616), 3313 (±423) gr among study groups respectively. Both parameters were significantly lower in the preeclampsia group compared to the control group (p<0.001, p<0.001). Gender of neonates did not differ significantly (p=0.548). Comparison of caesarean section rates revealed a statistically significant difference between the study groups, with the rate being significantly higher in the preeclampsia group compared with controls (94% vs. 60%, p=0.001). Apgar scores of the neonates were significantly lower in the preeclampsia group compared with the control group (p=0.001).

Comparisons of the laboratory and fetal Doppler parameters between the preeclampsia and control groups are presented in Table 2. Among fetal doppler parameters, the umbilical artery and mean uterine artery pulsatility indices were significantly higher in preeclampsia compared with control group (p=0.001 for both), whereas the middle cerebral artery pulsatility index did not differ significantly between study groups (p=0.479). Since uteroplacental insufficiency rate was higher in preeclamptic patients compared with control group, estimated fetal weight (EFW) of the fetuses was significantly lower in preeclampsia group (p=0.001). Mean IFN α-1 levels in the preeclamptic patients and control groups were 80.5 (±31.60) and 57.9 (±23.32) pmol/L respectively. Comparisons of IFN α-1 levels revealed a statistically significant difference between the study groups (p=0.001).

Patients diagnosed with early-onset preeclampsia were further divided into two groups -fetal growth restriction (FGR) group, non-FGR group- based on the occurrence of the FGR. Demographic, laboratory and fetal doppler parameters of the patients with early onset preeclampsia were presented in Table 3. No statistically significant difference was observed in terms of mean arterial pressure (p=0.753). Although maternal IFN α-1 levels were higher in the FGR compared with non-FGR group (83.4±29.57 pmol/L, 77.4±34.02 pmol/L), this difference did not reach statistical significance (p=0.504). The median spot urine protein creatinine ratio (mg/mg) and 24-hour urine protein levels (mg/day) in FGR and non-FGR group were 2410, 945 mg/mg, and 2214, 889 mg/day respectively. Although a statistically significant difference was observed between the groups in terms of spot urine protein-to-creatinine ratio (p=0.010), no statistically significant difference was observed for 24-hour urine protein (p=0.091). Among fetal doppler parameters umbilical artery pulsatility index was significantly higher in FGR group compared with non-FGR group (p=0.001). On the other hand, no statistically significant differences were found in the pulsatility indices of the middle cerebral artery and the mean uterine artery (p=0.185, p=0.158). Since the rate of low birth weight was higher in the FGR group, the Apgar scores of neonates were significantly lower in the FGR group compared with the non-FGR group (p=0.004, p=0.006).

Correlation analyses performed in the preeclampsia group between the maternal birth outcomes, urine protein parameters, and maternal IFN α-1 protein levels are presented in Table 4. There was no statistically significant correlation between the maternal IFN α-1 levels and birth weight, gestational age at delivery, mean arterial pressure, spot urine protein-to-creatinine ratio, 24-hour

Table 2. Comparisons of IFN α-1 and fetal doppler parameters between the preeclampsia and control groups

	Preeclampsia group (n=50)	Control group (n=50)	Total (n=100)	p-values[a]
GA at blood sampling (weeks)[c]	29.6 (±0.87)	29.3 (±0.69)	29.5 (±0.79)	0.079
UA-PI[c]	1.28 (±0.25)	0.90 (±0.15)	1.09 (±0.28)	**0.001**
MCA-PI[c]	1.79 (±0.18)	1.77 (±0.21)	1.78 (±0.19)	0.479
Mean uterine artery PI[c]	1.38 (±0.41)	0.86 (±0.10)	1.12 (±0.40)	**0.001**
EFW percentile[b]	7 (1-56)	46 (13-96)	28 (1-96)	**0.001**
Maternal IFN α-1 (pmol/L)[c]	80.5 (±31.60)	57.9 (±23.32)	69.2 (±29.89)	**0.001**

UA-PI: Pulsatility index of umbilical artery, MCA-PI: Pulsatility index of middle cerebral artery, EFW: Estimated fetal weight, IFN α-1: Interferon-alpha 1, GA: Gestational age, [a]: Level of significance p<0.05, [b]: Data that were not normally distributed are expressed as median (minimum-maximum) and were compared using Mann-Whitney U test, [c]: Continuous variables distributed normally are expressed as mean and standard deviation and are compared using Student's t-test

Table 3. Demographic characteristics, IFN α-1 levels and fetal doppler parameters of the patients according to occurrence of fetal growth restriction

		FGR (n=26)	Non-FGR (n=24)	Total (n=50)	p-values[a]
Mean arterial pressure (mm/hg)[c]		104.1 (±8.17)	104.9 (±10.26)	104.4 (±9.14)	0.753
GA at blood draw (weeks)[c]		29.7 (±0.81)	29.4 (±0.91)	29.6 (±0.87)	0.233
UA-PI[c]		1.39 (±0.21)	1.16 (±0.24)	1.28 (±0.25)	**0.001**
MCA-PI[c]		1.83 (±0.18)	1.76 (±0.17)	1.79 (0.18)	0.185
Mean uterine artery PI[c]		1.46 (±0.35)	1.30 (±0.45)	1.38 (±0.41)	0.158
Spot urine protein-to-creatinine ratio (mmHg)[b]		2410 (281-13506)	945 (119-4936)	1433 (119-13506)	**0.010**
24-hour urine protein (mg/day)[b]		2214 (262-10596)	889 (315-5290)	1180 (262-10596)	0.091
Maternal serum IFN α-1 (pmol/L)[c]		83.4 (±29.57)	77.4 (±34.02)	80.5 (±31.6)	0.504
1-minute Apgar score[b]		7 (2-8)	8 (4-8)	7 (2-8)	**0.004**
5-minute Apgar score[b]		8 (3-9)	9 (7-9)	8 (3-9)	**0.006**
Birth method[d]	**SVB**	1 (3.8%)	2 (8.3%)	3 (6%)	0.602
	CS	25 (96.2%)	22 (91.7%)	47 (94%)	

GA: Gestational age, CS: Caesarean section, SVB: Spontaneous vaginal birth, UA-PI: Pulsatility index of umbilical artery, MCA-PI: Pulsatility index of middle cerebral artery, IFN α-1: Interferon-alpha 1, [a]: Level of significance p<0.05, [b]: Data that were not normally distributed are expressed as median (minimum-maximum) and were compared using Mann-Whitney U test, [c]: Continuous variables distributed normally are expressed as mean and standard deviation and are compared using Student's t-test, [d]: Categorical variables were presented as percentage (%) and count (n) and were compared using the Fisher's exact test

Table 4. Correlation analyses between maternal serum IFN α-1 levels and clinical parameters in patients included in preeclampsia group

IFN α-1	r	p[a]
GA at delivery [b]	0.216	0.133
Birth weight[b]	0.198	0.168
Mean arterial pressure[b]	-0.012	0.934
24-hour urine protein (mg/day)[c]	0.085	0.559
Spot urine protein-to-creatinine ratio (mg/mg)[c]	-0.075	0.605
Apgar 1. minute[c]	0.341	**0.015**
Apgar 5. minute[c]	0.273	0.055

GA: Gestational age, IFN α-1: Interferon alfa, [a]: Level of significance p<0.05, [b]: r= Pearson's correlation, [c]: r= Spearman's correlation

protein level, and Apgar 5th-minute of neonates (p>0.05). A statistically significant but weak correlation was observed only between the maternal IFN α-1 levels and Apgar 1st-minute of neonates (p=0.015, r=0.341).

ROC analysis of maternal IFN α-1 levels demonstrated limited ability to discriminate between preeclamptic and healthy pregnancies [area under the curve (AUC) =0.589; 95% confidence interval: 0.382-0.795] (Figure 1). Optimal cut-off for maternal IFN α-1 levels corresponding to highest Youden's index (0.251) was determined to be >80.15 pmol/L with a sensitivity of 61% and a specificity of 64%. These findings indicate that although IFN α-1 levels are

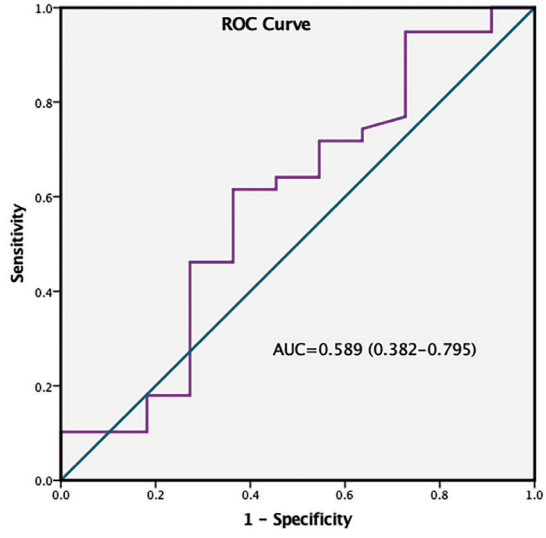

Figure 1. The receiver operating characteristic (ROC) curve analysis of maternal IFN α-1 levels for prediction of obstetric complications in patients diagnosed with preeclampsia

AUC: Area under the curve, IFN α-1: Interferon alpha-1

significantly elevated in preeclampsia, their discriminative performance is insufficient for clinical use as a standalone diagnostic biomarker.

Moreover, box-plot distributions and scatter-dot plot of maternal IFN α-1 levels in the preeclampsia and control groups are shown in Figure 2 and Figure 3.

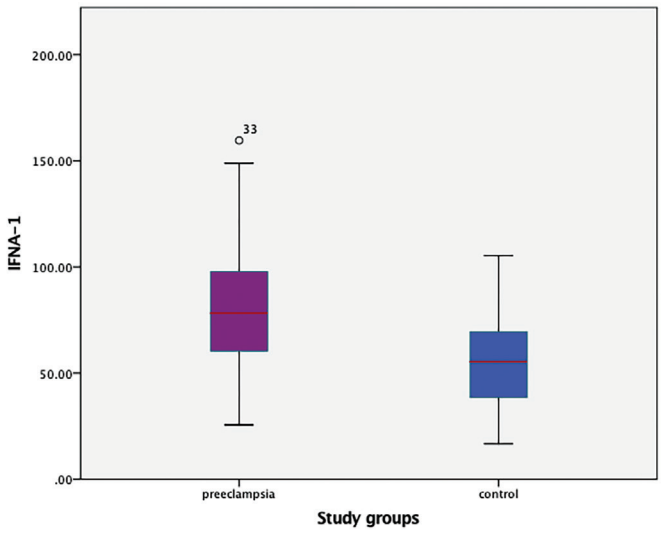

Figure 2. Box plot distribution of maternal IFN α-1 levels of preeclampsia and control groups

Box-plot distributions of maternal interferon alpha-1 (IFN α-1) levels in preeclampsia and control groups. The median is shown by the central line, boxes represent the 25th to 75th percentiles, whiskers indicate the range (excluding outliers)

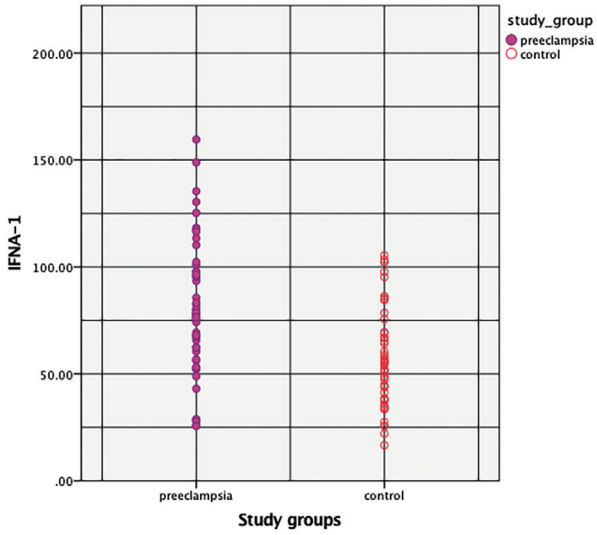

Figure 3. Scatter-dot distribution of maternal IFN α-1 levels in preeclampsia and control groups

Scatter-dot plot of maternal IFN α-1 levels in participants in the preeclampsia and control groups colour-coded by study groups; each dot represents an individual patient

IFN α-1: Interferon alpha-1

Discussion

In this study, we examined the role of maternal IFN α-1 in early-onset preeclampsia, a severe form of the disease that is closely linked to placental dysfunction. Our main finding is that maternal serum IFN α-1 levels were significantly higher in women with early-onset preeclampsia than in healthy, normotensive pregnant controls. This observation is consistent with previous reports indicating that increased type I interferon activity is associated with preeclampsia in high-risk autoimmune pregnancies and suggests that similar immunological mechanisms may also be relevant in early-onset preeclampsia outside autoimmune settings (12,13).

The relationship between type I interferons and pregnancy complications is not a new idea, but most previous research has focused on women with SLE or antiphospholipid syndrome, who already have high baseline IFN activity and a much greater risk of developing preeclampsia (12-14). Studies in these groups have shown that an increased type I *IFN* gene signature or higher IFN-α levels can appear before the clinical onset of preeclampsia, suggesting a possible causal role (13,14). Our results are consistent with these earlier findings but also expand their relevance, showing that IFN-α activation is present in early-onset preeclampsia even in women without known autoimmune disease. This indicates that an abnormal type I IFN response may be a more central part of preeclampsia pathophysiology than previously recognized.

Mechanistically, our findings support the idea that IFN-α plays a role in the main pathological features of preeclampsia: poor trophoblast invasion and widespread endothelial dysfunction. Recent experimental work using an “implantation-on-a-chip” model showed that exposure to type I IFN reduces the invasive ability of extravillous trophoblasts and produces a gene expression pattern similar to that seen in preeclampsia (16). This IFN-related trophoblast dysfunction may lead directly to shallow placentation, which is thought to trigger the disease process. In addition, IFN-α is a strong anti-angiogenic molecule. It can disturb the balance of angiogenic factors and, as suggested by Cerdeira and Thadhani (12), may make the maternal endothelium more sensitive to injury from circulating factors such as sFlt-1 (13). The higher umbilical and uterine artery pulsatility indices found in our preeclamptic group are in line with the severe placental insufficiency expected from these IFN-mediated mechanisms.

In the literature, there are also studies that have examined the relationship between IFN α-1 levels and vascular endothelial injury, abnormal placentation, preeclampsia, and adverse pregnancy outcomes. In the study by Hong et al. (17), which included pregnant women with SLE, those with higher plasma IFN levels were found to have a greater incidence of adverse pregnancy outcomes, including preeclampsia. In another study, Cappelletti et al. (18) demonstrated that increased type I IFN levels—thought to be secondary to heightened inflammation in preterm birth—were associated with an elevated type I IFN/IFN receptor ratio.

In another study conducted by Sacre et al. (19), type I IFN levels were observed to be more suppressed in SLE patients who were receiving hydroxychloroquine therapy. Furthermore, the findings of Seo et al. (20) showed that hydroxychloroquine treatment was associated with a reduced incidence of preeclampsia in pregnant women with SLE. Taken together, these two studies support the idea—similar to the implications of our own results—that type I IFN activity may play a meaningful role in the pathophysiology of preeclampsia. Similarly, Rahman et al. (21) reported that hydroxychloroquine significantly lowered TNF-α production and reduced endothelin-1 levels induced by preeclamptic serum. The drug also improved angiogenesis that had been impaired by TNF-α.

In another study conducted by Wada et al. (22), increased IFN-α levels were shown to exert an antiproliferative effect on human umbilical vein endothelial cells. These findings further support the possibility that elevated IFN-α may be associated with low birth weight and adverse pregnancy outcomes.

Overall, these studies indicate a strong link between higher IFN-α levels and vascular endothelial damage, abnormal placentation, and impaired vascularization. This relationship suggests that increased IFN-α activity may play a role in raising the risk of preeclampsia.

In addition to the immunological findings, pregnancies complicated by early-onset preeclampsia were associated with adverse perinatal outcomes, including lower gestational age at delivery and reduced birth weight, as expected. Apgar scores were also significantly lower in the preeclampsia group. The lower Apgar scores observed in the preeclampsia group are likely related to the higher rate of preterm delivery, which is a well-recognized consequence of early-onset preeclampsia.

Despite the clear increase in IFN α-1 levels in the preeclampsia group, our evaluation of its usefulness as an independent biomarker was not encouraging. The ROC analysis showed poor discriminatory performance (AUC=0.589), and the cut-off value of 80.15 pmol/L provided only moderate sensitivity (61%) and specificity (64%). This suggests that although IFN α-1 is associated with preeclampsia, the substantial overlap in values between cases and controls limits its value for clinical diagnosis or prediction.

This is an important finding because it indicates that IFN α-1 may reflect a broader inflammatory state rather than serving as a specific marker for preeclampsia itself. Given the complex and multifactorial nature of preeclampsia, it is unlikely that a single cytokine can fully represent the disease process. The better performance of biomarkers such as the sFlt-1/PlGF ratio, which capture the final common pathway of angiogenic imbalance, further highlights this complexity (1).

This study has several strengths, including its prospective case-control design, its focus on the clinically important and etiologically distinct early-onset preeclampsia phenotype, and the careful matching of control subjects. To our knowledge, it is the first study to specifically evaluate maternal serum IFN α-1 protein levels in a general (non-autoimmune) early-onset preeclampsia population and to assess its diagnostic value using ROC analysis.

Study Limitations

However, the study also has some limitations. Although the sample size was sufficient to detect differences in mean IFN α-1 levels, it may have been too small to identify more subtle associations, such as the non-significant trend toward higher IFN α-1 levels in the FGR subgroup. The cross-sectional design does not allow conclusions about causality or the temporal sequence between rising IFN α-1 levels and the onset of clinical symptoms. Measuring IFN α-1 at only a single time point may not reflect the dynamic changes in maternal immune function throughout pregnancy. In addition, we did not measure other cytokines or angiogenic markers, which would have enabled a more comprehensive assessment of the interactions between these biological pathways.

Conclusion

In conclusion, our study shows that maternal serum IFN α-1 levels are higher in women with early-onset preeclampsia, suggesting that changes in the type I interferon pathway

may play an important role in the disease. However, IFN α-1 alone does not seem to be a strong diagnostic marker. Future studies should follow patients over time to better understand when IFN-α levels start to rise in relation to the development of symptoms. Examining IFN α-1 together with other markers, such as angiogenic factors, may provide better predictive value. Although IFN α-1 by itself may not be sufficient for diagnosis, its association with preeclampsia indicates that the type I interferon pathway could be a useful target for new treatment strategies.

Ethics

Ethics Committee Approval: This prospective cross-sectional case-control study was conducted at the Perinatology Department of University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital. The study protocol was approved by the Local Ethics Committee (approval no. 2024-330, date: 10.09.2025), and all procedures were performed in accordance with the Declaration of Helsinki.

Informed Consent: Written informed consent was obtained from all patients after they were fully informed about the study procedures, potential risks and benefits, and their rights as participants.

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.B., V.A., Concept: B.B., F.E., Design: B.B., V.A., F.E., Data Collection or Processing: B.B., V.A., H.Ö.E., Analysis or Interpretation: B.B., H.Ö.E., Literature Search: B.B., F.E., H.Ö.E., Writing: B.B., V.A., H.Ö.E.

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

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

References

1. Dimitriadis E, Rolnik DL, Zhou W, Estrada-Gutierrez G, Koga K, Francisco RPV, et al. Pre-eclampsia. Nat Rev Dis Primers. 2023;9(1):8. Erratum in: Nat Rev Dis Primers. 2023;9(1):35.
2. Ives CW, Sinkey R, Rajapreyar I, Tita ATN, Oparil S. Preeclampsia-pathophysiology and clinical presentations: JACC state-of-the-art review. J Am Coll Cardiol. 2020;76(14):1690-1702.
3. Laresgoiti-Servitje E, Gómez-López N, Olson DM. An immunological insight into the origins of pre-eclampsia. Hum Reprod Update. 2010;16(5):510-524. Erratum in: Hum Reprod Update. 2011;17(2):291.
4. Raghupathy R. Cytokines as key players in the pathophysiology of preeclampsia. Med Princ Pract. 2013;22(Suppl 1): 8-19.
5. Denny MF, Thacker S, Mehta H, Somers EC, Dodick T, Barrat FJ, et al. Interferon-alpha promotes abnormal vasculogenesis in lupus: a potential pathway for premature atherosclerosis. Blood. 2007;110(8):2907-2915.
6. Thacker SG, Berthier CC, Mattinzoli D, Rastaldi MP, Kretzler M, Kaplan MJ. The detrimental effects of IFN-α on vasculogenesis in lupus are mediated by repression of IL-1 pathways: potential role in atherogenesis and renal vascular rarefaction. J Immunol. 2010;185(7):4457-4469.
7. Ashton SV, Whitley GS, Dash PR, Wareing M, Crocker IP, Baker PN, et al. Uterine spiral artery remodeling involves endothelial apoptosis induced by extravillous trophoblasts through Fas/FasL interactions. Arterioscler Thromb Vasc Biol. 2005;25(1):102-108.
8. Lyall F, Bulmer JN, Duffie E, Cousins F, Theriault A, Robson SC. Human trophoblast invasion and spiral artery transformation: the role of PECAM-1 in normal pregnancy, preeclampsia, and fetal growth restriction. Am J Pathol. 2001;158(5):1713-1721.
9. Roberts JM, Redman CW. Pre-eclampsia: more than pregnancy-induced hypertension. Lancet. 1993;341(8858):1447-1451.
10. Aldo PB, Krikun G, Visintin I, Lockwood C, Romero R, Mor G. A novel three-dimensional in vitro system to study trophoblast-endothelium cell interactions. Am J Reprod Immunol. 2007;58(2):98-110.
11. Kalkunte S, Lai Z, Tewari N, Chichester C, Romero R, Padbury J, et al. In vitro and in vivo evidence for lack of endovascular remodeling by third trimester trophoblasts. Placenta. 2008;29(10):871-878.
12. Cerdeira AS, Thadhani R. Editorial: "interfering" with preeclampsia. Arthritis Rheumatol. 2015;67(4):862-864.
13. Andrade D, Kim M, Blanco LP, Karumanchi SA, Koo GC, Redecha P, et al. Interferon-α and angiogenic dysregulation in pregnant lupus patients who develop preeclampsia. Arthritis Rheumatol. 2015;67(4):977-987.
14. Ugolini-Lopes MR, Torrezan GT, Gândara APR, Olivieri EHR, Nascimento IS, Okazaki E, et al. Enhanced type I interferon gene signature in primary antiphospholipid syndrome: association with earlier disease onset and preeclampsia. Autoimmun Rev. 2019;18(4):393-398.
15. Stockfelt M, Torell A, Gunnarsson I, Svenungsson E, Zickert A, Sennström MM, et al. Plasma interferon-alpha protein levels during pregnancy are associated with lower birth weight in systemic lupus erythematosus. Rheumatology (Oxford). 2025;64(3):1469-1475.
16. Simoni MK, Negatu SG, Park JY, Mani S, Arreguin MC, Amses KR, et al. Type I interferon exposure of an implantation-on-a-chip device alters invasive extravillous trophoblast function. Cell Rep Med. 2025;6(3):101991.
17. Hong S, Banchereau R, Maslow BL, Guerra MM, Cardenas J, Baisch J, et al. Longitudinal profiling of human blood transcriptome in healthy and lupus pregnancy. J Exp Med. 2019;216(5):1154-1169.
18. Cappelletti M, Presicce P, Lawson MJ, Chaturvedi V, Stankiewicz TE, Vanoni S, et al. Type I interferons regulate susceptibility to inflammation-induced preterm birth. JCI Insight. 2017;2(5):e91288.
19. Sacre K, Criswell LA, McCune JM. Hydroxychloroquine is associated with impaired interferon-alpha and tumor necrosis factor-alpha production by plasmacytoid dendritic cells in systemic lupus erythematosus. Arthritis Res Ther. 2012;14(3):R155.

20. Seo MR, Chae J, Kim YM, Cha HS, Choi SJ, Oh S, et al. Hydroxychloroquine treatment during pregnancy in lupus patients is associated with lower risk of preeclampsia. Lupus. 2019;28(6):722-730.

21. Rahman R, Murthi P, Singh H, Gurusinghe S, Mockler JC, Lim R, et al. The effects of hydroxychloroquine on endothelial dysfunction. Pregnancy Hypertens. 2016;6(4):259-262.

22. Wada H, Nagano H, Yamamoto H, Noda T, Murakami M, Kobayashi S, et al. Combination of interferon-alpha and 5-fluorouracil inhibits endothelial cell growth directly and by regulation of angiogenic factors released by tumor cells. BMC Cancer. 2009;9:361.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):307-318
DOI: 10.4274/BMB.galenos.2026.80774

Comparative Evaluation of Artificial Intelligence Models in Providing Patient-Oriented Information on Scabies: Implications for Public Health and Clinical Practice

Uyuz Hastalığı Hakkında Hasta Odaklı Bilgi Sağlamada Yapay Zeka Modellerinin Karşılaştırmalı Değerlendirmesi: Halk Sağlığı ve Klinik Uygulama Açısından Çıkarımlar

Orhan Şen

Kayseri City Hospital, Department of Dermatology, Kayseri, Turkey

Abstract

Objective: Scabies is an important public health problem, where patients have difficulty accessing information due to its high contagiousness and the stigma it causes. The aim of this study was to comparatively evaluate the accuracy, reliability, and readability of responses provided by current artificial intelligence (AI) chat models (ChatGPT, Google Gemini, and Microsoft Copilot) to frequently asked questions about scabies.

Method: In this cross-sectional study, conducted in January 2026, the 20 most frequently searched Turkish-language questions about scabies in Turkey were identified using Google Trends data. Responses generated by AI models were analyzed using the DISCERN overall quality rating, the global quality score (GQS), an accuracy score (1-6 Likert scale), and the Ateşman readability formula. Word counts were also compared. Statistical analyses were performed using the Friedman test for related samples, with Wilcoxon signed-rank post-hoc tests and Bonferroni correction.

Results: Google Gemini demonstrated significantly superior performance to ChatGPT and Microsoft Copilot across all evaluated scores: DISCERN overall quality rating (median: 5.0), GQS (median: 5.0), and accuracy (median: 6.0) ($p<0.05$). Microsoft Copilot received the lowest accuracy scores. In a readability analysis, Google Gemini

Öz

Amaç: Uyuz, hastaların yüksek bulaşıcılığı ve buna bağlı damgalanma nedeniyle bilgiye erişimde güçlük yaşadığı önemli bir halk sağlığı sorunudur. Bu çalışmanın amacı, uyuz hastalığı ile ilgili sık sorulan sorulara verilen yanıtlarda, güncel yapay zeka (YZ) sohbet modellerinin (ChatGPT, Google Gemini ve Microsoft Copilot) doğruluğunu, güvenilirliğini ve okunabilirliğini karşılaştırmalı olarak değerlendirmektir.

Yöntem: Ocak 2026 tarihinde gerçekleştirilen bu kesitsel çalışmada, Google Trends verileri kullanılarak Türkiye genelinde uyuz ile ilgili en sık aranan 20 Türkçe soru belirlenmiştir. YZ modelleri tarafından üretilen yanıtlar DISCERN genel kalite puanı, global kalite skoru (GQS), doğruluk skoru (1-6 Likert ölçeği) ve Ateşman okunabilirlik formülü ile analiz edilmiştir. Ayrıca yanıtların kelime sayıları karşılaştırılmıştır. İstatistiksel analizler, ilişkili örneklemler için Friedman testi ve Bonferroni düzeltmeli post-hoc Wilcoxon işaretli sıralar testi kullanılarak gerçekleştirilmiştir.

Bulgular: Google Gemini DISCERN genel kalite puanı (medyan: 5,0), GQS (medyan: 5,0) ve doğruluk (medyan: 6,0) puanlarının tamamında, ChatGPT ve Microsoft Copilot'a kıyasla istatistiksel olarak anlamlı düzeyde ($p<0,05$) daha üstün performans sergilemiştir. Microsoft Copilot, doğruluk açısından en düşük skorları almıştır.

Address for Correspondence: Orhan Şen, Kayseri City Hospital, Department of Dermatology, Kayseri, Turkey
E-mail: eru2017@gmail.com **ORCID:** orcid.org/0000-0002-8946-7330
Received: 13.01.2026 **Accepted:** 22.07.2026 **Epub:** 27.07.2026 **Publication Date:** 29.09.2026

Cite this article as: Şen O. Comparative evaluation of artificial intelligence models in providing patient-oriented information on scabies: implications for public health and clinical practice. Bagcilar Med Bull. 2026;11(3):307-318

Abstract

produced the most comprehensible texts (58.8 points, "moderate difficulty") and the most comprehensive responses, with a mean length of 297 words.

Conclusion: Google Gemini emerged as the model that provided the most accurate, reliable, and patient-friendly information on scabies. While recommending such tools as supplementary sources of information, dermatologists should emphasize the risk of AI hallucinations and the need for physician oversight.

Keywords: Artificial intelligence, ChatGPT, Google Gemini, health literacy, scabies

Öz

Okunabilirlik analizinde Google Gemini (58,8 puan) "orta güçlükte" ve en anlaşılır metinleri üretirken, aynı zamanda ortalama 297 kelime ile en kapsamlı yanıtları veren model olmuştur.

Sonuç: Google Gemini, uyuz hastalığı hakkında en doğru, güvenilir ve hasta dostu (okunabilir) bilgiyi sağlayan model olarak öne çıkmaktadır. Dermatologlar, hastalarına ek bilgi kaynağı olarak bu modelleri önerirken, YZ'nin "halüsinasyon" riskini ve hekim kontrolünün önemini vurgulamalıdır.

Anahtar kelimeler: ChatGPT, Google Gemini, sağlık okuryazarlığı, uyuz, yapay zeka

Introduction

Scabies is an ectoparasitic infestation caused by *Sarcoptes scabiei var. hominis*; it is characterized by intense pruritus, is highly contagious, and remains a common condition worldwide. According to the World Health Organization, approximately 200 million people are estimated to be affected globally at any given time (1). The disease may lead to dermatological manifestations, sleep disturbances, secondary bacterial infections, and a substantial deterioration in quality of life (2,3).

The contagious nature of scabies, together with widespread misconceptions within society, imposes a significant psychosocial burden on affected individuals and contributes to feelings of stigma (4). Consequently, patients may hesitate to seek medical care, experience embarrassment when reporting their symptoms to physicians, and ultimately delay diagnosis and treatment. Such delays in the management of this highly contagious disease increase the risk of transmission within households and the community, thereby posing a substantial public health concern.

In recent years, rapid advances in artificial intelligence (AI)–based large language models (LLMs) have initiated a new era in health information–seeking behavior (5). Chatbots offer continuous (24/7) accessibility and provide a safe environment in which patients can share symptoms they perceive as private or embarrassing without fear of judgment (6-8). As a result, individuals with suspected scabies frequently turn to AI-powered chatbots before seeking in-person medical consultation.

Contemporary AI-based conversational agents are capable of responding to symptom-related inquiries and providing information regarding disease prevention, clinical course, and available treatment options (9).

Numerous studies have examined the roles, benefits, and limitations of AI in healthcare. Although AI chat models continuously expand the scope of information they provide through regular updates, the accuracy and reliability of their outputs remain a matter of ongoing debate (7,8).

In addition to accuracy, the comprehensibility of chatbot-generated responses for patients is a critical consideration. ChatGPT, Google Gemini, and Microsoft Copilot currently rank among the top three AI chatbots in terms of active user numbers and accessibility (10). Given the increasing utilization of these platforms, systematic evaluation of their accuracy, quality, readability, and overall comprehensibility is essential.

To date, no study has specifically assessed the quality, reliability, accuracy, and readability of information generated by LLMs regarding scabies—a highly contagious dermatological disease with significant public health implications. Considering the growing reliance of patients on AI-based information sources prior to seeking professional medical advice, there is a notable gap in the existing literature.

Therefore, the present study aimed to comparatively evaluate the accuracy, quality, reliability, and comprehensibility of responses generated by three widely used AI chatbots—ChatGPT, Google Gemini, and Microsoft Copilot—to the 20 most frequently searched patient questions related to scabies, as identified through search engine queries (11).

Materials and Methods

Study Design and Data Collection

This cross-sectional observational study was conducted on January 8, 2026.

Ethics Statement

This study analyzed only publicly available search trend data from Google Trends and text outputs generated by publicly available AI chatbots. As the research did not involve any human participants, patient records, identifiable personal data, biological specimens, or animal subjects, formal ethics committee approval was not required. The study was conducted in compliance with the ethical principles of the Declaration of Helsinki (revised 2013).

Question Selection Protocol

To identify questions that best reflect public information-seeking behavior related to scabies, a structured, reproducible query-selection protocol was applied to the Google Trends database (Google LLC, Mountain View, CA, USA). The following predefined parameters were used for the search:

- Search term: "Uyuz" (the Turkish lay term for scabies)
- Geographic region: Turkey
- Time range: Last 5 years (January 2021-January 2026)
- Category filter: "Health"
- Search type: Web search
- Date of data extraction: January 8, 2026.

Within these parameters, the "Related Queries" section of Google Trends was examined, focusing on the "Top" subcategory, which lists the most frequently searched terms and question patterns associated with the main keyword. From the resulting list, queries were extracted in descending order of search frequency. The selection process was refined using the following predefined inclusion and exclusion criteria:

Inclusion Criteria

- Queries phrased as a question or directly representing a patient-oriented information need (e.g., transmission, symptoms, treatment, hygiene, prognosis).
- Queries written in Turkish.
- Queries clearly related to human scabies.

Exclusion Criteria:

- Duplicate queries or queries with overlapping semantic content (only the most frequently searched version was retained).
- Queries unrelated to the medical condition (e.g., idiomatic, slang, or metaphorical uses of the term "uyuz" in Turkish, which can also denote "annoying" or "unpleasant" in colloquial usage).
- Queries referring to veterinary scabies or non-medical contexts.
- Brand-, product-, or location-specific queries (e.g., pharmacy names, specific drug brands).

After applying these predefined inclusion criteria, the 20 highest-ranked queries were retained as the final question set. To enhance objectivity and minimize selection bias, the inclusion and exclusion criteria were defined before screening commenced, and only the search frequency ranking provided by Google Trends—not the author's subjective judgment of clinical relevance—was used to order the final list. The complete list of the 20 selected questions, grouped into thematic clinical categories, is provided in Table 1.

A total of 20 questions were selected to ensure sufficient content diversity while maintaining the feasibility of manual scoring and adequately addressing patients' primary concerns (Table 1).

Standardization of Question Wording

Because Google Trends displays related queries as short keyword strings rather than complete sentences, each retained query was converted into a single, grammatically complete Turkish question before submission. This standardization was limited to (i) rephrasing the keyword string as a full interrogative sentence, and (ii) correcting spelling and restoring Turkish diacritics omitted in the search string [Where two closely related top-ranked queries addressed the same information need (e.g., diagnosis and the need for a blood test), they were combined into a single question]. No clinical information or terminology absent from the original queries was added. The original Google Trends query strings, the standardized Turkish questions submitted verbatim to the AI models, and their English translations are provided in Supplementary Table S1.

AI Models and Standardized Querying Protocol

The selected 20 questions were submitted on the same day (January 8, 2026) to the three most up-to-date, publicly accessible AI models at that time:

1. ChatGPT (GPT-5.1 model, OpenAI, San Francisco, CA, USA),

2. Google Gemini (Gemini 3.0 model, Google LLC, Mountain View, CA, USA),

Table 1. The 20 most frequently asked questions regarding scabies identified via Google Trends data

No	Category	Question
1	Definition	What is scabies and what causes it?
2	Transmission	How is scabies transmitted from person to person?
3	Transmission	Is being in the same environment as someone with scabies enough for transmission?
4	Symptoms	What are the first symptoms of scabies and when do they start?
5	Symptoms	Why does scabies itching become more severe at night?
6	Symptoms	In which body parts are scabies rashes more commonly seen?
7	Risk groups	Can scabies be transmitted from animals (cats, dogs) to humans?
8	Risk groups	Is scabies seen only in people with poor personal hygiene?
9	Diagnosis	How is scabies diagnosed? Is a blood test required?
10	Treatment	What is the medical treatment for scabies? Which creams or pills are used?
11	Treatment	Is it normal for itching to continue after treatment is applied?
12	Treatment	How is scabies treatment performed in pregnant women and infants?
13	Treatment	Can scabies be treated with natural home remedies (vinegar, sulfur, etc.)?
14	Prevention	If one family member has scabies, should the others also be treated?
15	Hygiene	How should clothes and bed linens be washed to be cleared of scabies?
16	Hygiene	How should non-washable items (carpets, sofas, shoes) be cleaned?
17	Contagiousness	How soon can one return to work or school after starting treatment?
18	Contagiousness	How long can the scabies mite survive outside the human body (on objects)?
19	Complications	What serious health problems does scabies cause if left untreated?
20	Immunity	Can a person who has had scabies once get it again?

Note: The 20 questions represent the most frequently searched scabies-related queries identified through Google Trends (Turkey, Health category, January 2021-January 2026, "Top" related queries) and are grouped into thematic clinical categories. English translations are provided for presentation only; the original Google Trends query strings and the standardized Turkish questions submitted verbatim to all three AI models are listed in Supplementary Table S1.

3. Microsoft Copilot (GPT-5-based model, Microsoft Corp., Redmond, WA, USA).

A standardized querying protocol was applied to minimize variability and ensure comparability across models. Each of the 20 questions was entered verbatim, in identical Turkish wording, into all three AI models without any modification, paraphrasing, or reordering. No system prompts, custom instructions, persona settings, or hidden contextual prompts were used. All queries were submitted through the publicly accessible, default web interface of each model, and all default parameters were retained (no manual adjustment of temperature, top-p, response length, or any other configurable model parameter, as these are not user-modifiable in the standard consumer interfaces of the evaluated platforms). To minimize potential bias arising from prior user data, conversational memory, or personalization features, browser cookies and cache were cleared before each session; all queries were submitted in incognito or private browsing mode while the user was logged out of any associated account; and a new, independent chat session ("New Chat") was initiated for each individual question for each model, ensuring that no prior context could influence subsequent responses. For each question, only the first generated response was recorded and analyzed, and the "regenerate" function was not used to preserve the real-world user experience that patients would typically encounter when consulting these chatbots.

A total of 60 responses (20 questions ×3 models) were collected, saved in raw text format, anonymized, and numbered for blinded analysis.

Evaluation of AI-generated Responses

The 60 AI-generated responses were independently evaluated by two board-certified dermatologists, each with a minimum of five years of clinical experience in dermatology. In cases of disagreement, a third expert opinion was sought. To reduce assessment bias, the reviewers were blinded to the identity of the AI model that generated each response.

The following assessment tools were used:

DISCERN Instrument

The DISCERN instrument is a validated tool developed to assess the quality and reliability of written consumer health information. It comprises 15 items addressing reliability and the presentation of treatment information, followed by a final item (item 16) that provides an overall quality rating on a five-point scale (1: low quality with serious or extensive

shortcomings; 3: moderate quality with potentially important but not serious shortcomings; 5: high quality with minimal shortcomings) (12). In the present study, only this overall quality rating (item 16) was applied to each response, and the full DISCERN total score was not calculated. This measure is therefore referred to as the DISCERN overall quality rating throughout the manuscript.

Global Quality Score (GQS)

The GQS is a five-point scale used to evaluate the overall flow, accessibility, and usefulness of the content for users. A score of 1 represents "poor quality/inadequate flow", whereas a score of 5 indicates "excellent quality" (13).

Accuracy Score

A six-point Likert scale was employed to assess concordance of the AI-generated responses with current medical literature and dermatology guidelines (1: Completely incorrect; 2: Predominantly incorrect; 3: Equal amounts of correct and incorrect information; 4: More correct than incorrect; 5: Nearly completely correct; 6: Completely correct) (14).

To ensure transparency and reproducibility of the accuracy assessment, the following predefined reference standards were used as the gold standard to judge AI-generated responses:

- The 2020 International Alliance for the Control of Scabies (IACS) Consensus Criteria for the diagnosis of scabies (1).
- The 2017 European guideline for the management of scabies, jointly issued by the European Academy of Dermatology and Venereology (EADV) and the International Union against Sexually Transmitted Infections (IUSTI) (15).
- The Turkish Society of Dermatology (*Türk Dermatoloji Derneği*) "Scabies Diagnosis and Treatment Guideline", a Delphi-based national consensus document developed by experts in dermatology, parasitology, entomology, pediatrics, pharmacology, and public health, and adopted as the principal national reference for the diagnosis and management of scabies in Turkey (16).

For each response, the following content domains were systematically evaluated against the above reference standards: (i) Etiology and causative organism; (ii) modes of transmission; (iii) clinical features and symptom characteristics; (iv) anatomical distribution of lesions; (v) differential diagnosis; (vi) first-line and alternative pharmacological treatments, with appropriate dosing and duration; (vii) management of close contacts and household members; (viii) environmental decontamination measures; (ix) post-scabietic itching and expected clinical course; and (x) appropriate referral and red-flag indicators.

A response was considered fully correct only when all clinically relevant statements were consistent with the predefined reference standards listed above and contained no factual errors, outdated information, or potentially misleading recommendations. Statements that were factually inaccurate, that contradicted current guidelines, or that had the potential to mislead patients (e.g., endorsement of unproven home remedies as definitive treatment) were rated as incorrect, with the final score reflecting the overall proportion of correct versus incorrect content in the response.

Readability and Text Analysis

Word counts of all responses were recorded. The readability of Turkish texts was evaluated using the Ateşman readability formula, which generates a readability score ranging from 0 to 100 based on average word length (number of syllables) and sentence length (number of words) (17).

The Ateşman readability score was calculated using the following formula:

Ateşman score =198.825-[40.175 × average word length (syllables)]-[2.610 × average sentence length (words)]

The resulting scores were classified according to the following reference ranges:

90-100: Very easy

70-89: Easy

50-69: Moderately difficult

30-49: Difficult

0-29: Very difficult.

Statistical Analysis

Statistical analyses were performed using the Statistical Package for the Social Sciences (SPSS), version 26.0 (IBM Corp., Armonk, NY, USA). The normality of the data distribution was assessed using the Shapiro-Wilk test and histogram analysis. Numerical variables that did not follow a normal distribution were presented as the median and interquartile range (IQR).

Since the same 20 standardized questions were submitted to all three AI models, the resulting observations represented related (paired or repeated) measurements rather than independent samples. Therefore, differences in DISCERN overall quality ratings, GQS, accuracy scores, word counts,

and Ateşman readability scores among the three AI models were compared using the Friedman test, which is the appropriate non-parametric method for related samples. When the Friedman test indicated a statistically significant difference ($p<0.05$), pairwise comparisons were performed using the Wilcoxon signed-rank test. To account for multiple comparisons, the Bonferroni correction was applied, and pairwise differences were considered statistically significant at an adjusted threshold of $\alpha=0.05/3\approx0.017$; unadjusted p-values are reported for all pairwise comparisons.

The associations between response length (word count) and the DISCERN overall quality rating and accuracy score were examined using Spearman's rank correlation coefficient (rS), given the ordinal nature of the rating scales. For this analysis, the responses of all three AI models were pooled (n=60; 20 questions × 3 models).

Inter-rater reliability between the two independent evaluators was assessed using the intraclass correlation coefficient (ICC), and results were reported with 95% confidence intervals (CIs). For the Friedman tests, correlation analyses, and ICC, a two-sided p-value of <0.05 was considered statistically significant.

Results

Quality and Accuracy Analysis

When the DISCERN overall quality ratings of the AI models were evaluated, Google Gemini demonstrated the highest performance with a median score of 5.0 (IQR: 4.75-5.00). ChatGPT followed with a median score of 4.0 (IQR: 4.00-4.00), while Microsoft Copilot exhibited the lowest performance, with a median score of 4.0 (IQR: 3.00-4.00) (Table 2). The Friedman test revealed a statistically significant difference among the three AI models [$\chi^2(2)=21.58$, $p<0.001$].

Post-hoc pairwise comparisons using Wilcoxon signed-rank tests (unadjusted p-values; Bonferroni-adjusted significance threshold $\alpha=0.017$) showed that the DISCERN overall quality ratings of Google Gemini were significantly higher than those of both ChatGPT (p=0.008) and Microsoft Copilot (p=0.001). No statistically significant difference was observed between ChatGPT and Copilot (p=0.098).

A similar pattern was observed for the GQS, which evaluates content flow and usefulness. Google Gemini achieved a significantly higher median GQS score of 5.0 (IQR: 4.75-5.00) than ChatGPT (median: 4.0, IQR: 4.00-4.00) and Microsoft

Table 2. Comparison of DISCERN overall quality ratings, global quality scores (GQS), accuracy scores, and Ateşman readability scores of AI models

Variables	Model	Mean ± SD	Median (IQR)	Min-max	p-value†
DISCERN overall quality rating	ChatGPT	4.10±0.55	4.00 (4.00-4.00)	3-5	**<0.001**
	Google Gemini	**4.70±0.57**	**5.00 (4.75-5.00)**	**3-5**	
	Microsoft Copilot	3.70±0.47	4.00 (3.00-4.00)	3-4	
GQS score	ChatGPT	4.10±0.55	4.00 (4.00-4.00)	3-5	**<0.001**
	Google Gemini	**4.70±0.57**	**5.00 (4.75-5.00)**	**3-5**	
	Microsoft Copilot	3.70±0.47	4.00 (3.00-4.00)	3-4	
Accuracy score	ChatGPT	5.10±0.64	5.00 (5.00-5.25)	4-6	**<0.001**
	Google Gemini	**5.60±0.68**	**6.00 (5.00-6.00)**	**4-6**	
	Microsoft Copilot	4.45±0.60	4.50 (4.00-5.00)	3-5	
Word count	ChatGPT	172.6±37.6	171.0 (146.0-190.5)	118-283	**<0.001**
	Google Gemini	**297.4±28.3**	**293.5 (279.0-308.0)**	**254-352**	
	Microsoft Copilot	195.1±38.5	192.0 (159.8-220.3)	140-262	
Readability (Ateşman score)	ChatGPT	42.0±17.6	43.2 (40.0-49.7)	9.2-69.6	**<0.001**
	Google Gemini	**58.8±4.4**	**58.0 (56.1-59.9)**	**51.5-69.5**	
	Microsoft Copilot	52.2±13.5	56.3 (48.5-60.5)	2.8-63.3	

Ateşman readability score: Scores range from 0 to 100; higher scores indicate increased readability and clarity (0-29: Very difficult; 30-49: Difficult; 50-69: Moderately difficult; 70-89: Easy; 90-100: Very easy). †: Statistical analysis: The Friedman test was used for comparisons among related samples, followed by post-hoc Wilcoxon signed-rank tests with Bonferroni correction. p-values refer to the Friedman test ($p<0.05$ considered statistically significant); for post-hoc pairwise comparisons, the Bonferroni-adjusted significance threshold was $\alpha=0.017$. SD: Standard deviation, IQR: Interquartile range (Q1-Q3). Boldface: Indicates the values representing the statistically best performance in each respective category.

Copilot (median: 4.0, IQR: 3.00-4.00). The Friedman test demonstrated a statistically significant difference among the models [$\chi^2(2) = 21.58$, $p<0.001$], with post-hoc Wilcoxon comparisons ($\alpha=0.017$) confirming Gemini's superiority over both ChatGPT ($p=0.008$) and Copilot ($p=0.001$).

Inter-rater reliability between the two independent reviewers was excellent, with an ICC of 0.87 (95% CI: 0.82-0.91, $p<0.001$).

Comparison of Accuracy Scores

When evaluated for medical accuracy, Google Gemini achieved the highest median score of 6.0 (IQR: 5.00-6.00), corresponding to responses closest to being completely correct. ChatGPT ranked second (median score 5.0, IQR: 5.00-5.25), whereas Microsoft Copilot had the lowest accuracy (median score 4.5, IQR: 4.00-5.00).

The Friedman test revealed a statistically significant difference in accuracy scores among the models [$\chi^2(2) = 22.73$, $p<0.001$]. Post-hoc Wilcoxon signed-rank tests (unadjusted p-values evaluated against the Bonferroni-adjusted threshold of $\alpha=0.017$) indicated that all inter-model differences were statistically significant.

Gemini > ChatGPT ($p=0.012$)

ChatGPT > Copilot ($p=0.015$)

Gemini > Copilot ($p=0.001$).

These findings indicate that in the context of scabies-related information, Google Gemini outperformed the other models in both content quality and medical accuracy (Figure 1).

Readability and Response Length Analysis

The comprehensibility of texts generated by AI models was assessed using the Ateşman readability formula, and the level of detail was evaluated by word count.

Response Length (Word Count)

Google Gemini produced the most detailed and comprehensive responses, with a mean length of 297.4±28.3 words. This was followed by Microsoft Copilot with 195.1±38.5 words and ChatGPT with 172.6±37.6 words. The differences in response length among the models were statistically significant ($p<0.001$). On average, Gemini's responses were approximately 72% longer than ChatGPT's and 52% longer than Copilot's (Figure 2).

Readability

According to the Ateşman readability formula, Google Gemini (mean score: 58.8) and Microsoft Copilot (mean score: 52.2) generated texts classified as "moderately difficult", whereas ChatGPT (mean score: 42.0) produced more complex sentence structures and fell within the "difficult category" (Figure 2). Despite generating the longest responses, Google Gemini maintained the highest level of readability among the models ($p<0.001$).

Correlation Analysis

Spearman correlation analysis of the pooled responses (n=60) demonstrated positive associations between word count and both the DISCERN overall quality rating ($rS=0.40$, $p=0.002$) and the accuracy score ($rS=0.32$, $p=0.013$). These findings suggest that more detailed explanations are generally associated with higher information quality and accuracy. Scores ($r_S=0.32$, $p<0.05$) increased. These findings suggest that more detailed explanations are generally associated with higher medical quality and accuracy.

Discussion

To the best of our knowledge, this study is among the first to comparatively evaluate the performance of contemporary AI models in the context of scabies, a disease of critical public health importance in which patients frequently face barriers to accessing reliable information due to fear of stigma. The most striking finding of our study is that Google Gemini was statistically significantly superior to ChatGPT and Microsoft Copilot across all evaluated parameters, including the DISCERN overall quality rating, GQS (content flow), accuracy, and readability.

Our findings indicate substantial performance differences among AI models, which may have meaningful clinical and social implications.

Comparison of DISCERN Overall Quality Ratings and GQS

In our study, the Google Gemini model achieved significantly higher DISCERN overall quality ratings and GQS scores than ChatGPT and Microsoft Copilot. The DISCERN instrument is an internationally validated tool designed to assess the reliability of patient information materials and the balanced presentation of treatment options. Therefore, higher DISCERN overall quality ratings may reflect not only the quantity of information provided but also its structural integrity, neutrality, and trustworthiness (12).

Previous studies evaluating the quality of information generated by LLMs in dermatology have generally highlighted ChatGPT as a leading model. For example, Shapiro et al. (18) reported that ChatGPT achieved "moderate to good" DISCERN and GQS scores for general medical topics. Similarly, dermatology-focused studies have shown that ChatGPT provides fluent and well-structured responses, although it does not always present treatment alternatives in a balanced manner (19). In contrast, the higher DISCERN overall quality ratings and GQS scores observed for Google Gemini in our study are noteworthy.

Several factors may explain this difference. First, Gemini's stronger integration with up-to-date and comprehensive academic data sources within the Google ecosystem may have enabled more guideline-based and systematic responses. Second, Gemini's performance in Turkish natural language processing appeared superior, particularly patient-friendly phrasing and logical flow. This advantage

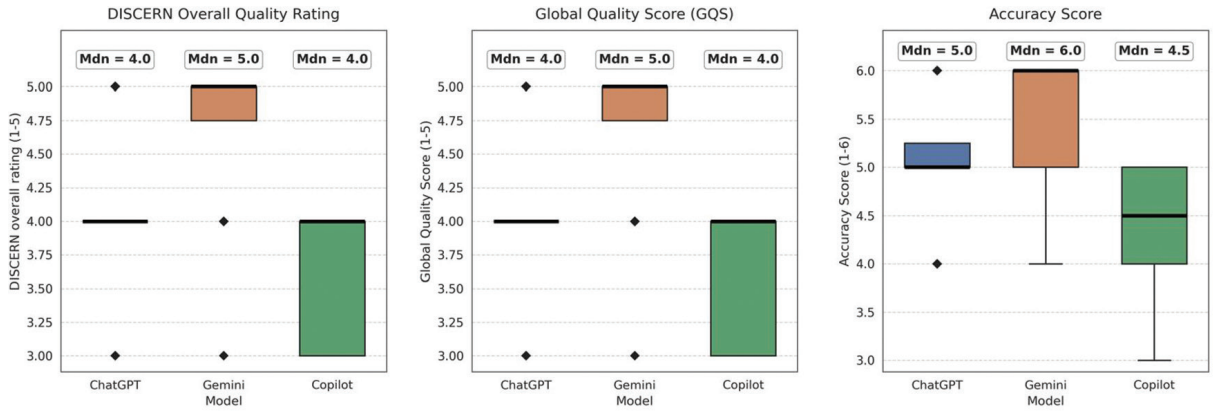

Figure 1. Distribution of DISCERN overall quality ratings, global quality scores (GQS), and accuracy scores of the artificial intelligence models (n=20 responses per model). Boxes represent the interquartile range (first to third quartile), and the horizontal line within each box indicates the median; when the median coincides with the first or third quartile, the median line overlaps the corresponding box edge, and a box collapsed to a single line indicates that the first quartile, median, and third quartile are identical. Median values are displayed above each box. Whiskers extend to the most extreme values within 1.5 × the interquartile range from the box edges (Tukey's method), and individual points beyond the whiskers represent outlying responses. Statistical comparisons were performed using the Friedman test for related samples, followed by post-hoc Wilcoxon signed-rank tests with Bonferroni correction

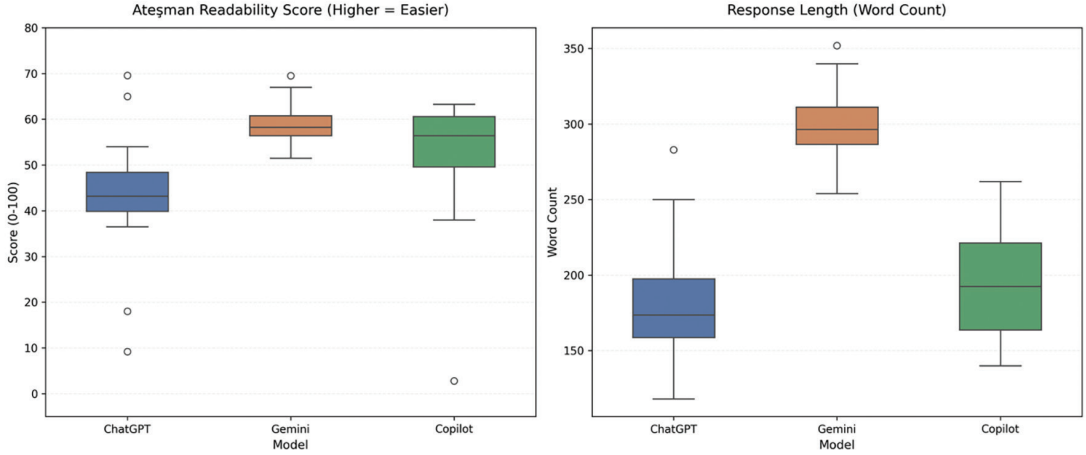

Figure 2. Ateşman readability scores and response lengths (word count) of the artificial intelligence models (n=20 responses per model). Boxes represent the interquartile range (first to third quartile), and the horizontal line within each box indicates the median. Whiskers extend to the most extreme values within 1.5 × the interquartile range from the box edges (Tukey's method), and open circles beyond the whiskers represent outlying responses. Higher Ateşman scores indicate easier readability. Statistical comparisons were performed using the Friedman test for related samples, followed by post-hoc Wilcoxon signed-rank tests with Bonferroni correction

was reflected in higher scores on scales assessing content flow and user utility, such as the GQS and the Ateşman readability score (20-22).

These findings contribute to the literature by demonstrating that LLM performance may vary depending on both disease context and language setting.

Comparison of Accuracy Scores

Accuracy analysis is one of the most clinically critical aspects of our study. When accuracy scores were examined, Google Gemini (median: 6.0) produced responses that were nearly error-free, whereas Microsoft Copilot (median: 4.5) provided information that was occasionally incomplete or potentially misleading. Our results indicate that Google Gemini achieved the highest accuracy in responses related to scabies, whereas Microsoft Copilot demonstrated lower accuracy.

In particular, misinformation regarding critical topics such as transmission routes and treatment protocols may contribute to increased disease spread, suggesting that Copilot should be used with greater caution in this context. ChatGPT, by contrast, demonstrated a "safe middle-ground" profile—less optimal than Gemini but more reliable than Copilot. Previous studies have reported variability in the accuracy performance of LLMs (23). For instance, in the study by Zorlu et al. (24) on hidradenitis suppurativa, Microsoft Copilot achieved relatively high accuracy scores, whereas in our study it exhibited the lowest accuracy. In another investigation by Valentini et al. (25) involving sarcoma patients, ChatGPT accuracy rates ranged between 70% and 85%, although some responses were found to be clinically incomplete or misleading.

This discrepancy may be attributed to Copilot's reliance on internet-based information retrieval (e.g., Bing Search), which may direct the model toward non-scientific forum or blog sources in conditions such as scabies, where misinformation and disinformation are prevalent. In contrast, Gemini appears to produce more guideline-oriented and reliable responses, likely because of its integration with Google's extensive academic data infrastructure.

In a comprehensive review on the use of AI in medicine, Topol emphasized that misinformation in infectious diseases can have serious public health consequences (26). Within this framework, the lower accuracy scores observed for Copilot in our study warrant particular caution, especially for diseases like scabies in which incorrect information is widespread in the community.

The relatively low accuracy of Copilot observed in our study represents a significant risk for common diseases that are prone to misinformation on forums, blogs, and social media. Inaccurate or incomplete information may contribute to increased transmission and delays in appropriate treatment (27).

Taken together, our findings suggest that Gemini's more guideline-consistent and coherent responses, in contrast to Copilot's occasional provision of incomplete or misleading information, underscore that the clinical and public health impact of LLM use—particularly in infectious diseases—may be highly dependent on model selection.

Readability and Response Length Analysis

In patient education, comprehensibility is as critical as the accuracy of the information provided (28). In our study, readability was assessed using the Ateşman readability formula, and Google Gemini was found to generate responses that were both the longest and the most readable. In contrast, ChatGPT produced shorter responses with more complex sentence structures, resulting in lower readability scores.

In a study evaluating online health information, McInnes and Haglund (29) reported that most patient education materials are written above the average literacy level of the general population. In our study, Gemini's use of clearer and more structured language suggests a more patient-friendly approach that contrasts with this general trend.

Campos et al. (30), in their analysis of AI-generated responses, noted that while models may produce grammatically fluent outputs, the use of complex sentence structures can adversely affect readability. This observation aligns with our findings, in which ChatGPT generated shorter yet less readable responses.

In the study by Zorlu et al. (24) on hidradenitis suppurativa, a positive relationship was identified between response length and quality scores; however, the authors emphasized that excessively technical language may hinder patient comprehension. This finding provides an important point of comparison for our results, as Google Gemini was able to maintain higher readability despite producing longer responses.

The positive correlation observed in our study between word count and both the DISCERN overall quality rating and the accuracy score suggests that more detailed explanations are generally associated with higher-quality, more accurate medical information. However, this finding

should be interpreted with caution and not as a simple "longer is better" rule. Response length is best understood as a proxy for explanatory depth and topical coverage rather than as a direct determinant of content quality. Increased response length has several potential drawbacks that, beyond a certain threshold, may compromise rather than enhance the value of patient-oriented information. (i) Excessive verbosity may introduce redundancy, diluting the most clinically relevant messages. (ii) Longer texts may reduce overall clarity by burying key recommendations within lengthy explanations. (iii) Higher word counts typically correlate with more complex sentence structures, which can reduce readability and impede comprehension in patients with limited health literacy. (iv) Extended responses may include peripheral or tangential content that, while not factually incorrect, distracts from the core informational need. Therefore, response length should be considered alongside, rather than as a substitute for, validated indicators of structural quality, accuracy, and readability.

In our study, the apparent advantage of longer responses depended critically on whether the additional length translated into substantively informative content delivered in patient-accessible language. Gemini's ability to preserve high readability despite producing the longest responses indicates that this model achieves a favorable balance between informational depth and comprehensibility—a balance not observed in models that produce shorter but linguistically more complex outputs (e.g., ChatGPT) or longer outputs without commensurate gains in clarity. This pattern reinforces the view that, in patient education contexts, the quality of information delivery is determined not by volume alone but by the integrated combination of accuracy, structural completeness, and readability (31).

Collectively, these findings underscore that when LLMs are used for patient education, evaluation should consider not only the amount of information provided but also how that information is presented.

Study Limitations

This study has several limitations that should be acknowledged when interpreting the findings. First, the evaluation was limited to a single dermatological condition (scabies). Although scabies is clinically and epidemiologically important and has substantial public health implications, the present findings cannot be directly extrapolated to other medical fields, dermatological disorders, or healthcare domains in which the performance of AI chatbots may differ depending on disease complexity, the volume and quality of available training data, and the degree to which authoritative guidelines are available. Disease-specific replication studies across diverse clinical areas are therefore warranted before drawing broader conclusions about the comparative utility of AI chatbots in patient education.

Second, the numbers of analyzed questions (n=20) and AI models (n=3) were inherently limited. Although these numbers were selected to balance content diversity with the feasibility of structured manual scoring, larger question sets and the inclusion of additional AI platforms in future studies would enhance external validity.

In addition, the correlation analysis pooled the responses of three models to the same 20 questions; these observations were therefore not fully independent, and the observed association may partly reflect between-model differences, as the model producing the longest responses also received the highest scores, rather than a within-model relationship between response length and quality.

Third, AI chatbots are subject to continuous, rapid updates—including changes to underlying model architecture, training data, fine-tuning procedures, and safety filters—that may meaningfully alter response quality, accuracy, and readability over time. Consequently, the findings reported here represent a cross-sectional snapshot of model performance as of January 8, 2026, and should not be interpreted as reflecting the long-term or future capabilities of any of the evaluated platforms. Periodic re-evaluation will be necessary to track the evolution of AI model performance in healthcare contexts.

Fourth, AI-generated outputs are inherently stochastic, and the same prompt may yield slightly different responses on repeated queries. In the present study, only the first generated response to each query was analyzed to reflect the real-world experience of a patient consulting these chatbots. While this design enhances ecological validity, it does not capture potential intra-model output variability. Future studies employing repeated queries with statistical pooling, or sensitivity analyses across paraphrased versions of the same prompt, could provide additional insight into the reproducibility of AI-generated medical information and the extent of prompt-sensitivity effects.

Fifth, the study was conducted exclusively in Turkish. Given that LLM performance—including factual accuracy, fluency, and readability—may vary substantially across languages and cultural contexts, the generalizability of

the present results to other linguistic settings is limited. Cross-linguistic studies would help clarify the extent to which the observed performance differences are language-dependent.

Finally, although the accuracy assessment was anchored to predefined international and national reference standards (the IACS Consensus Criteria, the EADV/IUSTI European guideline, and the Turkish Society of Dermatology guideline), expert evaluation inevitably involves a degree of subjectivity. The high inter-rater reliability observed in this study (ICC =0.87) supports the robustness of the assessment, but fully objective, automated evaluation frameworks may further enhance reproducibility in future research.

Moreover, only the overall quality item of DISCERN was applied rather than the full 16-item instrument, and because DISCERN was originally designed for information on treatment choices, this rating may capture the quality of non-treatment responses (e.g., those on transmission or hygiene) less precisely.

Conclusion

This study is among the first to perform a comparative evaluation of the performance of LLMs in delivering patient-oriented information about scabies in terms of quality, accuracy, and readability. Our findings indicate that although all evaluated models generally provided acceptable levels of information, significant differences were observed in content quality, medical accuracy, and information presentation.

Google Gemini achieved the highest performance in DISCERN overall quality ratings and GQS scores, as well as in accuracy assessments. ChatGPT yielded moderately consistent results, whereas Microsoft Copilot exhibited lower performance, particularly in critical areas such as treatment guidance and contact management. These findings indicate that LLMs do not provide consistent quality in patient education and that model selection plays a decisive role in the reliability of the information delivered.

Analyses of readability and response length revealed that more detailed and structured responses are generally associated with higher quality and accuracy scores. However, presenting information in a clear, simplified, and patient-appropriate manner was found to be as important as the quantity of information provided.

Large language models may serve as supportive tools for patient education in common and highly contagious dermatological conditions such as scabies. Nevertheless, patients should be clearly informed that AI-generated outputs do not replace professional medical advice. The use of AI-based information resources, within the framework of physician guidance and up-to-date clinical guidelines, is essential to ensure both patient safety and public health.

Ethics

Ethics Committee Approval: Because the study is based on the analysis of publicly available data and AI outputs, and does not involve human or animal subjects, it does not require ethical committee approval.

Informed Consent: Not necessary.

Footnotes

Acknowledgements

The author would like to thank Ilgaz Genç, MD, and Merve Cansu Kaya, MD, for their contribution as independent evaluators in the blinded assessment of the AI-generated responses.

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

Supplementary Table Link: https://d2v96fxpocvxx.cloudfront.net/688d2d00-d207-464d-89b6-73f393f4f50c/content-images/f8c97ac1-de33-4d2a-bf8c-c22793e00ca6.pdf

References

1. Engelman D, Yoshizumi J, Hay R, Osti M, Micali G, Norton S, et al. The 2020 international alliance for the control of scabies consensus criteria for the diagnosis of scabies. Br J Dermatol. 2020;183(5):808-820.
2. Karimkhani C, Colombara DV, Drucker AM, Norton SA, Hay R, Engelman D, et al. The global burden of scabies: a cross-sectional analysis from the global burden of disease study 2015. Lancet Infect Dis. 2017;17(12):1247-1254.
3. Romani L, Steer AC, Whitfeld MJ, Kaldor JM. Prevalence of scabies and impetigo worldwide: a systematic review. Lancet Infect Dis. 2015;15(8):960-967.
4. Worth C, Heukelbach J, Fengler G, Walter B, Liesenfeld O, Feldmeier H. Impaired quality of life in adults and children with scabies from an impoverished community in Brazil. Int J Dermatol. 2012;51(3):275-282.
5. Thirunavukarasu AJ, Ting DSJ, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. Nat Med. 2023;29(8):1930-1940.
6. Recker F, Neubauer R, Wittek A, Scholten N. Large language models and women's health: a digital companion for informed decision-making. Arch Gynecol Obstet. 2025;312(3):663-670.
7. Wang L, Wan Z, Ni C, Song Q, Li Y, Clayton E, et al. Applications and concerns of ChatGPT and other conversational large language

models in health care: systematic review. J Med Internet Res. 2024;26:e22769.

8. Chow JC, Li K. Large language models in medical chatbots: opportunities, challenges, and the need to address AI risks. Information. 2025;16(7):549.

9. Yun HS, Bickmore T. Online health information–seeking in the era of large language models: cross-sectional web-based survey study. J Med Internet Res. 2025;27:e68560.

10. Index.dev. ChatGPT Stats 2025: 800M users, traffic data & usage breakdown [Internet]. Index.dev; 2025. Available from: https://www.index.dev/blog/chatgpt-statistics

11. Google LLC. Google Trends. Mountain View, CA: Google LLC; [accessed: 08 January 2026]. Available from: https://trends.google.com

12. Charnock D, Shepperd S, Needham G, Gann R. DISCERN: an instrument for judging the quality of written consumer health information on treatment choices. J Epidemiol Community Health. 1999;53(2):105-111.

13. Yeung AWK. Evaluation of content quality of online health information by global quality score: a case study of researchers misnaming it and citing secondary sources. Publications. 2025;13(2):23.

14. Goodman R, Patrinely J, Stone Jr C, Zimmerman E, Donald R, Chang S, et al. Accuracy and reliability of chatbot responses to physician questions. JAMA Netw Open. 2023;6(10):e2336483.

15. Salavastru CM, Chosidow O, Boffa MJ, Janier M, Tiplica GS. European guideline for the management of scabies. J Eur Acad Dermatol Venereol. 2017;31(8):1248-1253.

16. Türk Dermatoloji Derneği. Uyuz tanı ve tedavi rehberi [İnternet]. Ankara: Türk Dermatoloji Derneği. Erişim adresi: https://turkdermatoloji.org.tr/icerik/detay/401

17. Ateşman E. Türkçede okunabilirliğin ölçülmesi. Dil Dergisi. 1997;58:71-74.

18. Shapiro J, Avitan-Hersh E, Greenfield B, Khamaysi Z, Dodiuk-Gad RP, Valdman-Grinshpoun Y, et al. The use of a ChatGPT-4-based chatbot in teledermatology: a retrospective exploratory study. JDDG J Dtsch Dermatol Ges. 2025;23(3):311-319.

19. Lewandowski M, Łukowicz P, Świetlik D, Barańska-Rybak W. ChatGPT-3.5 and ChatGPT-4 dermatological knowledge level based on the specialty certificate examination in dermatology. Clin Exp Dermatol. 2024;49(7):686-691.

20. Fattah FH, Salih AM, Salih AM, Asaad SK, Ghafour AK, Bapir R, et al. Comparative analysis of ChatGPT and Gemini (Bard) in medical inquiry: a scoping review. Front Digit Health. 2025;7:1482712.

21. Arbel Y, Gimmon Y, Shmueli L. Evaluating the potential of large language models for vestibular rehabilitation education: comparison of ChatGPT, Google Gemini, and clinicians. Phys Ther. 2025;105(4):pzaf010.

22. Sami MA, Samad MA, Parekh K, Suthar PP. Comparative accuracy of ChatGPT 4.0 and Google Gemini in answering pediatric radiology text-based questions. Cureus. 2024;16(10):e70897.

23. Kung TH, Cheatham M, Medenilla A, Sillos C, De Leon L, Elepaño C, et al. Performance of ChatGPT on USMLE: potential for AI assisted medical education using large language models. PLoS Digit Health. 2023;2(2):e0000198.

24. Zorlu Ö, Günal U, Yazici S. The quality, accuracy, and readability of information about hidradenitis suppurativa provided by artificial intelligence: comparative analysis of ChatGPT, Copilot, and Perplexity. Medicine (Baltimore). 2025;104(39):e44728.

25. Valentini M, Szkandera J, Smolle MA, Scheipl S, Leithner A, Andreou D. Artificial intelligence large language model ChatGPT: is it a trustworthy and reliable source of information for sarcoma patients? Front Public Health. 2024;12:1303319.

26. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.

27. Engelman D, Steer AC. Control strategies for scabies. Trop Med Infect Dis. 2018;3(3):98.

28. Weiss BD. Health literacy and patient safety: help patients understand. Manual for clinicians. American Medical Association Foundation; 2007.

29. McInnes N, Haglund BJ. Readability of online health information: implications for health literacy. Inform Health Soc Care. 2011;36(4):173-189.

30. Campos VM, Heringer DL, Costa GP, Oliva HN, Anand A. Readability, linguistic complexity, and stigma in ChatGPT responses to opioid use disorder FAQs: a comparative analysis. Am J Addict. 2026;35(3):362-369.

31. Li Y, Ouyang H, Lin G, Peng Y, Yao J, Chen Y. Evaluation of the measurement properties of online health information quality assessment tools: a systematic review. Int J Nurs Sci. 2025;12(2):130-136.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):319-326
DOI: 10.4274/BMB.galenos.2026.21548

Respiratory Rate at Admission as an Independent Predictor of Intensive Care Requirement in Pregnant Women with COVID-19: A Retrospective Cohort Study

Başvuru Sırasındaki Solunum Sayısının COVID-19'lu Gebe Kadınlarda Yoğun Bakım Gereksiniminin Bağımsız Bir Prediktörü Olarak Değerlendirilmesi: Retrospektif Kohort Çalışması

Karolin Ohanoğlu Çetinel[1], Gülçin Hilal Alay[2], Emilya Zeynalli[1], Tasnim Sabouni[1], Rümeysa Akpınar[1], İsmail Alay[3]

[1]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Obstetrics and Gynecology, İstanbul, Turkey
[2]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Anesthesia and Intensive Care, İstanbul, Turkey
[3]Medicana Bahçelievler Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey

Abstract

Objective: To compare demographic, clinical, laboratory, and obstetric characteristics of pregnant women hospitalized with coronavirus disease-2019 (COVID-19) according to intensive care unit (ICU) requirement and to identify independent predictors of ICU admission.

Method: This retrospective cohort study included pregnant women hospitalized with real-time reverse transcription polymerase chain reaction (RT-PCR)-confirmed COVID-19 at a tertiary referral center between September 2020 and December 2021. Patients were categorized into ICU and non-ICU groups based on clinical course. Demographic, clinical, laboratory, and obstetric parameters recorded at hospital admission were compared between groups. Multivariable logistic regression analysis was performed to determine independent predictors of ICU admission.

Results: A total of 402 pregnant women were included, of whom 49 (12.2%) required ICU admission. Women admitted to the ICU were older and presented at an earlier gestational age than non-ICU patients. At admission, ICU patients had significantly higher respiratory rates and lower oxygen saturation levels. Laboratory

Öz

Amaç: Koronavirüs hastalığı-2019 (COVID-19) nedeniyle hastaneye yatırılan gebelerin yoğun bakım ünitesi (YBÜ) gereksinimine göre demografik, klinik, laboratuvar ve obstetrik özelliklerini karşılaştırmak ve YBÜ yatışının bağımsız belirleyicilerini saptamaktır.

Yöntem: Bu retrospektif kohort çalışmaya, Eylül 2020-Aralık 2021 tarihleri arasında üçüncü basamak bir referans merkezinde gerçek zamanlı ters transkripsiyon polimeraz zincir reaksiyonu (RT-PCR) ile doğrulanmış COVID-19 tanısı ile hastaneye yatırılan gebeler dahil edildi. Hastalar klinik seyirlerine göre YBÜ ve YBÜ dışı gruplar olarak sınıflandırıldı. Başvuru anındaki demografik, klinik, laboratuvar ve obstetrik veriler gruplar arasında karşılaştırıldı. YBÜ yatışının bağımsız belirleyicilerini tanımlamak amacıyla çok değişkenli lojistik regresyon analizi yapıldı.

Bulgular: Toplam 402 gebe çalışmaya dahil edildi ve bunların 49'u (%12,2) YBÜ'ye yatırıldı. YBÜ'ye kabul edilen gebeler, YBÜ dışı hastalara kıyasla daha ileri yaşta olup daha erken gebelik haftasında başvurmuşlardı. Başvuru anında YBÜ hastalarında solunum sayısı anlamlı derecede daha yüksek, oksijen satürasyonu ise daha düşüktü. Laboratuvar analizlerinde, YBÜ hastalarında lenfosit sayıları anlamlı

Address for Correspondence: Karolin Ohanoğlu Çetinel, University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Obstetrics and Gynecology, İstanbul, Turkey
E-mail: karolette@gmail.com **ORCID:** orcid.org/0000-0002-3223-3293
Received: 23.01.2026 **Accepted:** 22.07.2026 **Epub:** 24.07.2026 **Publication Date:** 29.09.2026

Cite this article as: Ohanoğlu Çetinel K, Alay GH, Zeynalli E, Sabouni T, Akpınar R, Alay İ. Respiratory rate at admission as an independent predictor of intensive care requirement in pregnant women with COVID-19: a retrospective cohort study. Bagcilar Med Bull. 2026;11(3):319-326

Abstract

evaluation revealed significantly lower lymphocyte counts and markedly higher C-reactive protein and procalcitonin levels in the ICU group, whereas hemoglobin, hematocrit, total leukocyte count, neutrophil count, platelet count, D-dimer, and fibrinogen levels did not differ significantly between groups. Adverse obstetric outcomes were more frequent among ICU patients, including higher rates of cesarean delivery, preterm birth, and maternal mortality. Median hospital stay was significantly longer in ICU patients compared with non-ICU patients [18 (interquartile range, 13.5-23) vs. 5 (interquartile range, 3-8) days; p<0.001]. In multivariable analysis, respiratory rate at admission remained an independent predictor of ICU admission (adjusted odds ratio: 1.30 per 1 breath/min increase; 95% confidence interval: 1.20-1.41; p<0.001).

Conclusion: In pregnant women hospitalized with COVID-19, respiratory rate at admission was independently associated with ICU admission. Respiratory rate is a simple, readily available, and clinically meaningful marker that may facilitate early risk stratification and timely management in this population.

Keywords: COVID-19, hypoxemia, intensive care unit, pregnancy, respiratory failure

Öz

olarak daha düşük; C-reaktif protein ve prokalsitonin düzeyleri ise belirgin şekilde daha yüksekti. Hemoglobin, hematokrit, toplam lökosit sayısı, nötrofil sayısı, trombosit sayısı, D-dimer ve fibrinojen düzeyleri açısından gruplar arasında anlamlı fark saptanmadı. Obstetrik sonuçlar YBÜ hastalarında daha kötüydü; sezaryen doğum, preterm doğum ve maternal mortalite oranları daha yüksekti. Hastanede kalış süresinin medyanı YBÜ grubunda YBÜ dışı gruba göre anlamlı derecede daha uzundu [18 (çeyrekler arası aralık, 13,5-23) güne karşılık 5 (çeyrekler arası aralık, 3-8) gün; p<0,001]. Çok değişkenli analizde, başvuru anındaki solunum sayısı YBÜ yatışı için bağımsız ve güçlü bir öngördürücü olarak saptandı (düzeltilmiş odds oranı: 1,30; her 1 nefes/dak artış için; %95 güven aralığı: 1,20-1,41; p<0,001).

Sonuç: COVID-19 nedeniyle hastaneye yatırılan gebelerde, başvuru anındaki solunum sayısı YBÜ yatışı ile bağımsız olarak ilişkili bulunmuştur. Başvuru anındaki solunum sayısı, kolay erişilebilir ve klinik olarak anlamlı bir belirteç olup bu hasta grubunda erken risk sınıflandırmasına ve zamanında klinik yönetime katkı sağlayabilir.

Anahtar kelimeler: COVID-19, gebelik, hipoksemi, solunum yetmezliği, yoğun bakım ünitesi

Introduction

Pregnancy is characterized by profound immunologic, respiratory, and hemodynamic adaptations that, while essential for fetal development, may simultaneously increase maternal vulnerability to severe respiratory infections such as coronavirus disease-2019 (COVID-19) (1-3). These changes include a shift toward Th2-dominant immunity, reduced cell-mediated immune responses, and altered cytokine regulation, potentially weakening maternal antiviral defense. Concurrently, pregnancy is associated with progressive reductions in functional residual capacity due to diaphragmatic elevation, increased oxygen consumption, and progesterone-mediated increases in minute ventilation, resulting in reduced respiratory reserve and a predisposition to rapid oxygen desaturation during pulmonary infections (2,4-6).

Although most pregnant women with COVID-19 experience mild disease, accumulating evidence indicates that pregnancy itself is an independent risk factor for adverse outcomes, including pneumonia, need for oxygen therapy, intensive care unit (ICU) admission, and mechanical ventilation (7-9). Several large international cohorts have demonstrated significantly higher rates of maternal morbidity and mortality among pregnant women with severe acute respiratory syndrome-coronavirus-2 (SARS-CoV-2) infection compared with non-pregnant reproductive-age women. Comorbidities such as obesity, diabetes, chronic hypertension, and advanced maternal age further amplify susceptibility to respiratory failure via compounded inflammatory and metabolic dysregulation (10-12).

As the pandemic evolved, attention increasingly focused on the identification of early clinical and biochemical markers capable of predicting deterioration. Inflammatory markers—particularly C-reactive protein (CRP), ferritin, and D-dimer—were repeatedly associated with disease severity, reflecting the hyperinflammatory and prothrombotic state characteristic of severe COVID-19 (13,14). Respiratory parameters at presentation, including oxygen saturation and respiratory rate, have emerged as simple but powerful predictors of impending decompensation in adults; however, obstetric-specific data remain scarce, and the unique physiology of pregnancy may alter their predictive performance.

Despite extensive global research, pregnancy-specific risk stratification models applicable at the time of hospital admission remain limited, underscoring the need for simple and clinically applicable predictors. Most published cohorts include mixed populations or report limited respiratory or laboratory parameters; very few studies have incorporated comprehensive maternal, obstetric, and laboratory data simultaneously. For clinicians managing pregnant women with COVID-19, early identification of those at high risk for deterioration is crucial to guide triage decisions, allocate

monitoring resources, plan timely delivery when indicated, and coordinate care with anesthesiology and critical care teams.

Turkey, with its high-volume tertiary obstetric centers and centralized referral pathways, has cared for thousands of pregnant women with COVID-19 throughout the pandemic; however, detailed analyses of predictors of ICU admission in this population are lacking. The present study aimed to evaluate clinical, respiratory, inflammatory, and obstetric parameters associated with ICU admission in pregnant women hospitalized with COVID-19. In this context, clinically applicable pregnancy-specific risk stratification tools available at hospital admission remain limited, particularly those based on simple bedside parameters. By identifying early predictors measurable at presentation, this study seeks to contribute pregnancy-specific evidence to guide risk stratification and to support multidisciplinary maternal care during severe SARS-CoV-2 infection.

Materials and Methods

Study Design and Setting

This retrospective cohort study was conducted at a high-volume tertiary care referral center providing comprehensive obstetric and critical care services in İstanbul, Turkey. The study included pregnant women hospitalized with real-time reverse transcription polymerase chain reaction (RT-PCR)-confirmed COVID-19 infection between September 2020 and December 2021. The multidisciplinary care team, consisting of obstetricians, anesthesiologists, and intensive care specialists, jointly managed the patients and contributed to data acquisition.

Study Population

All pregnant women admitted with symptomatic COVID-19 during the study period were screened for eligibility. Patients were categorized into two groups based on clinical management:

1. ICU group: Women who required admission to the ICU due to respiratory deterioration or hemodynamic instability, and

2. Non-ICU group: Women who were managed on standard obstetric wards.

Exclusion criteria included missing essential clinical data, admission for non-COVID-related conditions, or unconfirmed SARS-CoV-2 infection. A total of 402 women met the inclusion criteria, of whom 49 (12.2%) required ICU care.

Data Collection

Clinical data were extracted from electronic medical records and included:

- Demographic and obstetric characteristics: Maternal age, gravida, parity, in vitro fertilization (IVF) conception, COVID-19 vaccination status, gestational age at admission, and pre-existing medical conditions (chronic disease, hypertensive disorders, gestational diabetes).

- Clinical findings at presentation: Blood pressure, heart rate, respiratory rate, and oxygen saturation (SpO_2).

- Laboratory parameters obtained at admission: Hemoglobin, hematocrit, white blood cell count, neutrophils, lymphocytes, platelets, D-dimer, fibrinogen, procalcitonin and CRP.

- Obstetric outcomes: Mode of delivery, gestational age at birth, ongoing pregnancy at discharge, and preterm birth (<37 weeks).

- Maternal outcomes: ICU admission, length of stay and maternal mortality. Length of stay refers to ICU stay for ICU patients and to the total hospital stay for non-ICU patients.

Laboratory tests were performed using standardized, automated analyzers in the hospital's central laboratory. Patients with missing key laboratory or respiratory parameters at admission were excluded.

Outcomes

The primary outcome was ICU admission due to COVID-19-related respiratory failure or clinical deterioration. Secondary outcomes included mode of delivery, preterm birth, maternal mortality, and laboratory differences between ICU and non-ICU patients.

Patients were admitted to the ICU according to predefined institutional criteria, including the need for high-flow nasal oxygen therapy, non-invasive or invasive mechanical ventilation, persistent hypoxemia (SpO_2 <93% despite supplemental oxygen), signs of increased work of breathing, hemodynamic instability requiring vasopressor support, or clinical deterioration unresponsive to standard ward-based management.

Ethics

The study was approved by the Institutional Review Board of the University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital (approval no: 149; date: 07/2021). The requirement for informed consent was waived due to the retrospective nature of the study. The study was conducted in accordance with the Declaration of Helsinki.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics (version 29.0). Continuous variables were assessed for normality using the Shapiro-Wilk test. Normally distributed continuous variables are presented as mean ± standard deviation and were compared using Student's t-test, while non-normally distributed variables are presented as median (interquartile range) and were compared using the Mann-Whitney U test. Categorical variables were summarized as frequencies and percentages and compared using the chi-square or Fisher's exact test, as appropriate. A multivariable logistic regression model was constructed to identify independent predictors of ICU admission, with variables selected based on clinical relevance and statistical significance in univariable analyses. Adjusted odds ratios (ORs) with 95% confidence intervals (CIs) were calculated. Model performance was assessed using the area under the receiver operating characteristic curve (AUC). Laboratory parameters were not included in the multivariable model to avoid overfitting and collinearity, given the limited number of ICU events. A two-sided p-value <0.05 was considered statistically significant.

Results

A total of 402 pregnant women with confirmed COVID-19 were included, of whom 49 (12.2%) required ICU admission. The mean maternal age was 29.6±5.7 years overall.

Women admitted to the ICU were significantly older than non-ICU patients (31.9±4.9 vs. 29.3±5.8 years, p=0.004). Gravidity and parity were comparable between the ICU and non-ICU groups. ICU patients also presented at a markedly earlier gestational age (29.7±5.7 vs. 33.0±7.9 weeks, p<0.001). Rates of chronic illness, hypertensive disorders, gestational diabetes, IVF pregnancies, and COVID-19 vaccination status did not differ significantly between groups (Table 1).

ICU patients presented with significantly more severe respiratory compromise than non-ICU patients. Respiratory rate was markedly elevated (32.0±10.2 vs. 19.1±4.6 breaths/min, p<0.001), and peripheral SpO_2 was significantly reduced (93.5±6.8% vs. 98.2±3.1%, p<0.001). These differences are shown in Figure 1 (respiratory rate distribution) and Figure 2 (oxygen saturation distribution), which both show clear separation between ICU and non-ICU groups. Blood pressure did not differ significantly between groups, whereas pulse rate was slightly higher among ICU patients (Table 2).

At hospital admission, patients requiring ICU care had significantly lower lymphocyte counts compared with non-ICU patients. Neutrophil and platelet counts did not differ significantly between groups. Inflammatory markers, including CRP and procalcitonin, were significantly higher among ICU patients, whereas hemoglobin, hematocrit, total leukocyte count, D-dimer, and fibrinogen levels were comparable between groups (Table 3).

Obstetric outcomes differed substantially between groups. Cesarean delivery occurred more frequently among ICU women (57.1% vs. 32.6%, p=0.001), whereas spontaneous

Table 1. Baseline demographic and clinical characteristics of pregnant women with COVID-19

Variable	ICU (n=49)	Non-ICU (n=353)	p-value
Age (years)	31.9±4.9	29.3±5.8	0.004
Gestational age at admission (weeks)	29.7±5.7	33.0±7.9	<0.001
Chronic disease (%)	18.3	11.9	0.21
Hypertensive disorder (%)	4.1	3.4	0.78
GDM (%)	10.2	7.9	0.54
IVF pregnancy (%)	4.1	6.2	0.56
COVID-19 vaccination (%)	12.2	17.6	0.41
Gravida, median (IQR)	2 (2-3.5)	2 (2-3)	0.63
Parity, median (IQR)	1 (1-2)	1 (0-2)	0.76

Values are presented as mean ± standard deviation, median (interquartile range), or n (%), as appropriate. Comparisons for continuous variables were performed using Student's t-test or Mann-Whitney U test, as appropriate. Categorical variables were compared using the chi-square or Fisher's exact test. ICU: Intensive care unit, GDM: Gestational diabetes mellitus, IVF: In vitro fertilization, COVID-19: Coronavirus disease-2019, IQR: Interquartile range

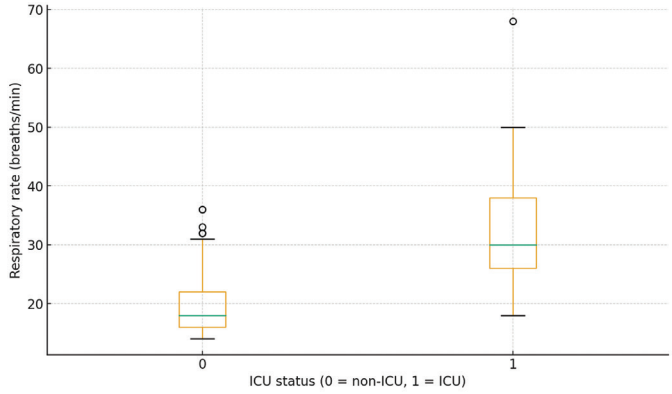

Figure 1. Respiratory rate according to ICU status. Boxplot showing respiratory rate distributions among ICU and non-ICU patients. ICU patients demonstrated significantly higher respiratory rates at admission

ICU: Intensive care unit

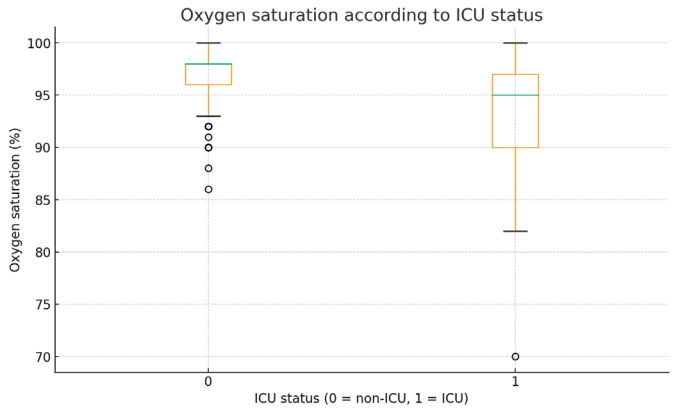

Figure 2. Peripheral oxygen saturation at admission according to ICU status. Boxplot illustrating significantly lower oxygen saturation in patients who required ICU care compared with those managed on the general ward

ICU: Intensive care unit

Length of hospital stay was substantially longer in ICU patients than in non-ICU patients [median 18 (IQR, 13.5-23) vs. 5 (IQR, 3-8) days, p<0.001] (Table 4).

In multivariable logistic regression analysis, respiratory rate at admission emerged as the strongest independent predictor of ICU admission among pregnant women hospitalized with COVID-19 (adjusted OR 1.30 per 1 breath/min increase; 95% CI 1.20-1.41; p<0.001) (Table 5).

After adjustment for maternal age and gestational age at admission, oxygen saturation was not independently associated with ICU admission. Maternal age showed a borderline association with ICU admission, whereas gestational age at admission was not significantly associated with ICU admission. In an exploratory ROC analysis, respiratory rate at admission demonstrated excellent discriminatory ability for predicting ICU requirement, yielding an AUC of 0.90 (Figure 3). Using Youden's index, the optimal respiratory rate cut-off for predicting ICU admission was identified as ≥25 breaths/min, providing a sensitivity of 79.6% and a specificity of 86.5%. This threshold offers clinically meaningful discrimination for early risk stratification on hospital admission.

Discussion

In this retrospective cohort study of pregnant women hospitalized with COVID-19, we evaluated clinical, respiratory, laboratory, and obstetric factors associated with ICU admission. The principal finding of our study was that respiratory rate at admission emerged as the strongest

Table 2. Clinical characteristics and vital signs at admission according to ICU requirement

Variable	ICU (n=49)	Non-ICU (n=353)	p-value
Oxygen saturation (%), mean ± SD	93.5±6.8	98.2±3.1	<0.001
Respiratory rate (breaths/min), mean ± SD	32.0±10.2	19.1±4.6	<0.001
Pulse rate (beats/min), mean ± SD	103±16	96±14	0.001
Systolic blood pressure (mmHg), mean ± SD	118±14	116±13	0.33
Diastolic blood pressure (mmHg), mean ± SD	72±10	71±9	0.48

Values are given as mean ± standard deviation. Respiratory rate data were available for 49 ICU and 333 non-ICU patients. SpO_2: Peripheral oxygen saturation, SD: Standard deviation, BP: Blood pressure, ICU: Intensive care unit

Table 3. Laboratory parameters at admission in ICU and non-ICU patients

	ICU (n=49) median (IQR)	Non-ICU (n=353) median (IQR)	p-value
Hemoglobin (g/dL)	11.05 (10.70-12.07)	11.40 (10.55-12.40)	0.402
Hematocrit (%)	33.20 (31.55-36.00)	34.80 (32.15-37.20)	0.092
White blood cell count (×10⁹/L)	8.49 (5.84-10.87)	8.27 (6.37-10.31)	0.776
Neutrophils (×10⁹/L)	6.71 (4.77-8.49)	6.18 (4.62-7.76)	0.255
Lymphocytes (×10⁹/L)	1.01 (0.81-1.38)	1.45 (0.98-1.88)	0.021
Platelets (×10⁹/L)	194.00 (167.00-234.00)	201.00 (163.00-246.50)	0.774
C-reactive protein (mg/L)	105.75 (50.38-134.00)	27.70 (7.97-59.80)	<0.001
D-dimer (mg/L)	1.42 (0.96-1.91)	1.25 (0.87-2.08)	0.699
Procalcitonin (ng/mL)	0.15 (0.10-0.28)	0.08 (0.05-0.14)	0.002
Fibrinogen (mg/dL)	525.00 (445.00-619.00)	495.50 (426.25-573.00)	0.086

Values are presented as median (interquartile range). Group comparisons were performed using the Mann-Whitney U test. ICU: Intensive care unit, IQR: Interquartile range

Table 4. Obstetric and maternal outcomes according to ICU requirement

Variable	ICU	Non-ICU	p-value
Cesarean delivery, n/N (%)	31/35 (88.6)	121/199 (60.8)	0.001
Vaginal delivery, n/N (%)	4/35 (11.4)	78/199 (39.2)	0.001
Ongoing pregnancy at discharge, n/N (%)	11/49 (22.4)	122/353 (34.6)	0.091
Preterm birth (<37 weeks), n/N (%)	24/27 (88.9)	42/177 (23.7)	<0.001
Gestational age at delivery (weeks), mean ± SD	32.0±4.1 (n=27)	37.6±3.2 (n=177)	<0.001
Length of stay (days), median (IQR)	18 (13.5-23), n=47	5 (3-8), n=332	<0.001
Maternal mortality, n/N (%)	14/49 (28.6)	0/353 (0)	<0.001

Data are presented as n/N (%), mean ± standard deviation (SD), or median [interquartile range (IQR)], as appropriate. Delivery mode was evaluated among women with a documented cesarean or vaginal delivery (ICU, n=35; non-ICU, n=199). Preterm birth and gestational age at delivery were evaluated among women who delivered and had available gestational age data (ICU, n=27; non-ICU, n=177). Length-of-stay data were available for 47 ICU and 332 non-ICU patients. ICU: Intensive care unit

Table 5. Multivariable logistic regression analysis for ICU admission

	Adjusted OR	95% CI	p-value
Oxygen saturation (per 1% increase)	1.00	0.88-1.14	0.951
Respiratory rate (per 1 breath/min increase)	1.30	1.20-1.41	<0.001
Maternal age (per 1 year increase)	1.08	1.00-1.17	0.063
Gestational age at admission (per 1 week increase)	1.01	0.96-1.07	0.677

Gestational age recorded as weeks and days was parsed (W and D) and modeled as weeks + days/7 in regression analyses. Adjusted ORs with 95% CIs are shown. ICU: Intensive care unit, CI: Confidence interval, OR: Odds ratio

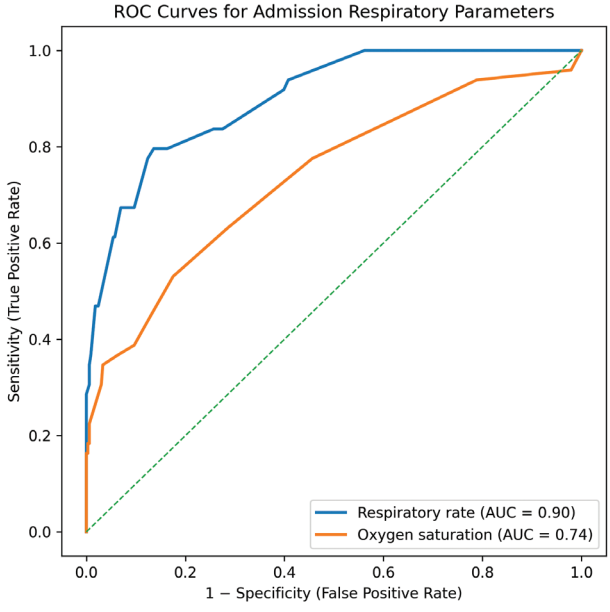

Figure 3. ROC curve illustrating the discriminatory performance of respiratory rate and oxygen saturation measured at hospital admission for predicting ICU requirement. Respiratory rate demonstrated excellent discrimination (AUC=0.90). An optimal cut-off value of 25 breaths/min provided 79.6% sensitivity and 86.5% specificity for predicting ICU admission, while oxygen saturation showed moderate discrimination (AUC=0.74)

AUC: Area under the curve, ROC: Receiver operating characteristic, ICU: Intensive care unit

independent predictor of ICU requirement, whereas oxygen saturation, maternal age, and gestational age did not retain independent significance after multivariable adjustment. This finding highlights the importance of early bedside respiratory assessment in pregnant patients with SARS-CoV-2 infection (15).

Consistent with previous studies, pregnant women requiring ICU care presented with more severe respiratory compromise at hospital admission, including higher respiratory rates and lower oxygen saturation levels (16,17). However, in the adjusted model, oxygen saturation was no longer independently associated with ICU admission, while respiratory rate remained a robust predictor. This observation is clinically plausible, as respiratory rate reflects increased work of breathing and early respiratory distress, often preceding overt hypoxemia. In pregnancy, physiological respiratory adaptations may further limit the discriminatory value of oxygen saturation alone, rendering respiratory rate a more sensitive indicator of impending clinical deterioration (18).

Inflammatory biomarkers, including CRP and procalcitonin, were significantly elevated among ICU patients in univariable analyses, consistent with previous reports linking systemic inflammation to severe COVID-19 (16,19,20). However, because laboratory parameters were not included in the multivariable model, their independent

association with ICU admission could not be assessed in the present study. Although vaccination status did not differ significantly between ICU and non-ICU groups, overall vaccination rates in the cohort were low, reflecting limited vaccine availability during the early phases of the pandemic. This likely reduced the ability to detect a protective effect of vaccination in this study.

Gestational age at admission was lower among ICU patients in unadjusted analyses; however, this association did not persist in multivariable analyses. Similarly, maternal age demonstrated only a borderline association with ICU admission. These findings indicate that acute respiratory compromise, rather than baseline maternal or obstetric characteristics, is the principal determinant of ICU requirement in pregnant women with COVID-19 (17,18).

Adverse obstetric outcomes, including higher rates of cesarean delivery and preterm birth, were more frequent among women admitted to the ICU, in line with previous reports (15,16). These outcomes are likely driven by maternal clinical deterioration and iatrogenic delivery decisions for maternal indications rather than a direct effect of SARS-CoV-2 infection on pregnancy itself (4). The higher rate of preterm birth observed in ICU patients likely reflects iatrogenic delivery decisions driven by maternal clinical deterioration rather than spontaneous preterm labor, underscoring the indirect impact of severe maternal COVID-19 on obstetric outcomes.

Furthermore, the observed maternal mortality rate among ICU patients was higher than that in many contemporary cohorts and likely reflected the severity of respiratory failure, delayed presentation, and the predominance of early pandemic variants prior to widespread vaccination and standardized treatment protocols.

The strengths of this study include a relatively large cohort from a tertiary referral center, a well-defined ICU population, and a comprehensive evaluation of clinical and respiratory parameters using multivariable analysis. Through an emphasis on readily available bedside measurements, our findings offer direct clinical applicability for early risk stratification in pregnant women with COVID-19.

Study Limitations

This study has several limitations. Its retrospective design may have introduced selection bias. Clinical management strategies for COVID-19 evolved substantially during the study period; these changes included the introduction of corticosteroids, anticoagulation protocols, and antiviral therapies. In addition, changing viral variants and evolving treatment protocols over time could not be fully accounted for in the analysis, and therefore may have influenced clinical outcomes. Furthermore, length of stay was defined as ICU stay for critically ill patients and as total hospital stay for non-ICU patients, which may limit direct comparison of hospitalization duration between groups.

A respiratory rate of 25 breaths/min or higher at presentation represents a practical bedside threshold that identifies pregnant patients at high risk for clinical deterioration, supporting its integration into early triage algorithms for COVID-19 in pregnancy.

Conclusion

Our findings demonstrate that respiratory rate at admission is a powerful and independent predictor of ICU admission among pregnant women hospitalized with COVID-19. Incorporating respiratory rate into early triage algorithms may improve timely identification of pregnant patients at risk for critical illness.

Ethics

Ethics Committee Approval: The study was approved by the Institutional Review Board of the University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital (approval no: 149; date: 07/2021).

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

Footnotes

Authorship Contributions

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

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

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

References

1. Rasmussen SA, Smulian JC, Lednicky JA, Wen TS, Jamieson DJ. Coronavirus disease 2019 (COVID-19) and pregnancy: what obstetricians need to know. Am J Obstet Gynecol. 2020;222(5):415-426.

2. Goodnight WH, Soper DE. Pneumonia in pregnancy. Crit Care Med. 2005;33(10 Suppl):S390-S397.

3. Kourtis AP, Read JS, Jamieson DJ. Pregnancy and infection. N Engl J Med. 2014;370(23):2211-2218.

4. Allotey J, Stallings E, Bonet M, Yap M, Chatterjee S, Kew T, et al.; for PregCOV-19 Living Systematic Review Consortium. Clinical manifestations, risk factors, and maternal and perinatal outcomes of coronavirus disease 2019 in pregnancy: living systematic review and meta-analysis. BMJ. 2020;370:m3320. Update in: BMJ. 2022;377:o1205.

5. Villar J, Ariff S, Gunier RB, Thiruvengadam R, Rauch S, Kholin A, et al. Maternal and neonatal morbidity and mortality among pregnant women with and without COVID-19 infection: the INTERCOVID multinational cohort study. JAMA Pediatr. 2021;175(8):817-826. Erratum in: JAMA Pediatr. 2022;176(1):104.

6. La Verde M, Riemma G, Torella M, Cianci S, Savoia F, Licciardi F, et al. Maternal death related to COVID-19: a systematic review and meta-analysis focused on maternal co-morbidities and clinical characteristics. Int J Gynaecol Obstet. 2021;154(2):212-219.

7. Wei SQ, Bilodeau-Bertrand M, Liu S, Auger N. The impact of COVID-19 on pregnancy outcomes: a systematic review and meta-analysis. CMAJ. 2021;193(16):E540-E548.

8. Della Gatta AN, Rizzo R, Pilu G, Simonazzi G. Coronavirus disease 2019 during pregnancy: a systematic review of reported cases. Am J Obstet Gynecol. 2020;223(1):36-41.

9. Lokken EM, Huebner EM, Taylor GG, Hendrickson S, Vanderhoeven J, Kachikis A, et al.; Washington State COVID-19 in Pregnancy Collaborative. Disease severity, pregnancy outcomes, and maternal deaths among pregnant patients with severe acute respiratory syndrome coronavirus 2 infection in Washington State. Am J Obstet Gynecol. 2021;225(1):77.e1-77.e14.

10. Di Toro F, Gjoka M, Di Lorenzo G, De Santo D, De Seta F, Maso G, et al. Impact of COVID-19 on maternal and neonatal outcomes: a systematic review and meta-analysis. Clin Microbiol Infect. 2021;27(1):36-46.

11. Chen L, Li Q, Zheng D, Jiang H, Wei Y, Zou L, et al. Clinical characteristics of pregnant women with COVID-19 in Wuhan, China. N Engl J Med. 2020;382(25):e100.

12. Adhikari EH, Moreno W, Zofkie AC, MacDonald L, McIntire DD, Collins RRJ, et al. Pregnancy outcomes among women with and without severe acute respiratory syndrome coronavirus 2 infection. JAMA Netw Open. 2020;3(11):e2029256.

13. Mendoza M, Garcia-Ruiz I, Maiz N, Rodo C, Garcia-Manau P, Serrano B, et al. Pre-eclampsia-like syndrome induced by severe COVID-19: a prospective observational study. BJOG. 2020;127(11):1374-1380.

14. Zambrano LD, Ellington S, Strid P, Galang RR, Oduyebo T, Tong VT, et al; CDC COVID-19 Response Pregnancy and Infant Linked Outcomes Team. Update: characteristics of symptomatic women of reproductive age with laboratory-confirmed SARS-CoV-2 infection by pregnancy status - United States, January 22-October 3, 2020. MMWR Morb Mortal Wkly Rep. 2020;69(44):1641-1647.

15. Metz TD, Clifton RG, Hughes BL, Sandoval G, Saade GR, Grobman WA, et al; for the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) Maternal-Fetal Medicine Units (MFMU) Network. Disease severity and perinatal outcomes of pregnant patients with coronavirus disease 2019 (COVID-19). Obstet Gynecol. 2021;137(4):571-580.

16. Knight M, Bunch K, Vousden N, Morris E, Simpson N, Gale C, et al; UK Obstetric Surveillance System SARS-CoV-2 Infection in Pregnancy Collaborative Group. Characteristics and outcomes of pregnant women admitted to hospital with confirmed SARS-CoV-2 infection in UK: national population based cohort study. BMJ. 2020;369:m2107.

17. Lapinsky SE. Management of acute respiratory failure in pregnancy. Semin Respir Crit Care Med. 2017;38(2):201-207.

18. Liu H, Wang LL, Zhao SJ, Kwak-Kim J, Mor G, Liao AH. Why are pregnant women susceptible to COVID-19? An immunological viewpoint. J Reprod Immunol. 2020;139:103122.

19. Henry BM, de Oliveira MHS, Benoit S, Plebani M, Lippi G. Hematologic, biochemical and immune biomarker abnormalities associated with severe illness and mortality in coronavirus disease 2019 (COVID-19): a meta-analysis. Clin Chem Lab Med. 2020;58(7):1021-1028.

20. Guan WJ, Ni ZY, Hu Y, Liang WH, Ou CQ, He JX, et al; China Medical Treatment Expert Group for Covid-19. Clinical characteristics of coronavirus disease 2019 in China. N Engl J Med. 2020;382(18):1708-1720.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):327-332
DOI: 10.4274/BMB.galenos.2026.45822

Association Between Obesity Severity and Inflammatory Markers in Obese Children

Obez Çocuklarda Obezite Şiddeti ile Enflamatuvar Belirteçler Arasındaki İlişki

Aysun Boğa, Ezgi Nafile Sayman, Sümeyra Gedik Çalışkan, Feyza Nur Akan Çelen, Elif Betül Keskin Çetinkaya, Yeşim Coşkun, Şirin Güven

University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, Department of Pediatrics, İstanbul, Turkey

Abstract

Objective: Childhood obesity is associated with chronic low-grade inflammation. This study aimed to evaluate the association between obesity severity and inflammatory hematologic parameters in children with obesity.

Method: This retrospective cross-sectional study included 88 pediatric patients with obesity aged 5-17 years who presented to the outpatient clinics of a tertiary pediatric training and research hospital. Patients were classified into obese and severely obese groups according to body mass index standard deviation score. Hematologic inflammatory parameters including neutrophil-to-lymphocyte ratio (NLR), platelet-to-lymphocyte ratio, monocyte-to-lymphocyte ratio (MLR) and hemoglobin levels were evaluated and compared between groups.

Results: A significantly greater proportion of participants were female in the severely obese group (p=0.016), whereas hemoglobin concentrations were higher among those with obesity (p=0.026). Among inflammatory hematologic ratios, only the MLR was significantly elevated in the obese group (p=0.003). In age- and sex-based subgroup analyses, an unadjusted difference in NLR was observed only in boys aged 5-9 years (p=0.016); however, this difference did not remain statistically significant after Bonferroni correction (adjusted p=0.064).

Conclusion: Inflammatory hematologic markers did not show a consistent relationship with obesity severity in children. Although MLR was lower in children with severe obesity and remained

Öz

Amaç: Çocukluk çağı obezitesi kronik düşük dereceli enflamasyon ile ilişkilidir. Bu çalışmada, obezite şiddeti ile obez çocuklardaki enflamatuvar hematolojik parametreler arasındaki ilişkinin değerlendirilmesi amaçlanmıştır.

Yöntem: Bu retrospektif kesitsel çalışmaya, üçüncü basamak bir çocuk eğitim ve araştırma hastanesinin polikliniklerine başvuran, yaşları 5-17 arasında değişen obezitesi olan 88 pediatrik hasta dahil edildi. Hastalar beden kitle indeksi standart sapma skoruna göre obez ve ağır obez gruplarına ayrıldı. Nötrofil-lenfosit oranı (NLR), trombosit-lenfosit oranı, monosit-lenfosit oranı (MLR) ve hemoglobin düzeylerini içeren hematolojik enflamatuvar parametreler değerlendirildi ve gruplar arasında karşılaştırıldı.

Bulgular: Ağır obez grubunda kadın cinsiyet oranı anlamlı olarak daha yüksek bulundu (p=0,016), buna karşın hemoglobin düzeyleri obez grubunda daha yüksekti (p=0,026). Enflamatuvar hematolojik oranlar arasında yalnızca MLR obez grupta anlamlı olarak daha yüksekti (p=0,003). Yaş ve cinsiyete göre yapılan alt grup analizlerinde, yalnızca 5-9 yaş arası erkek çocuklarda NLR açısından düzeltilmemiş analizde farklılık saptandı (p=0,016); ancak bu farklılık Bonferroni düzeltmesi sonrasında istatistiksel anlamlılığını korumadı (düzeltilmiş p=0,064).

Sonuç: Enflamatuvar hematolojik belirteçler çocuklarda obezite şiddeti ile tutarlı bir ilişki göstermemiştir. MLR ağır obez çocuklarda daha düşük bulunmuş ve yaş ile cinsiyete göre düzeltme sonrasında da obezite şiddeti ile anlamlı ilişkisini korumuşken, NLR çoklu karşılaştırma düzeltmesi sonrasında anlamlı bir ilişki göstermemiştir.

Address for Correspondence: Aysun Boğa, University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, Department of Pediatrics, İstanbul, Turkey
E-mail: aysunboga@gmail.com **ORCID:** orcid.org/0000-0002-9574-9529
Received: 28.02.2026 **Accepted:** 07.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Boğa A, Nafile Sayman E, Gedik Çalışkan S, Akan Çelen FN, Keskin Çetinkaya EB, Coşkun Y, et al. Association between obesity severity and inflammatory markers in obese children. Bagcilar Med Bull. 2026;11(3):327-332

Abstract

significantly associated with obesity severity after adjustment for age and sex, NLR showed no significant association after correction for multiple comparisons. These findings suggest that inflammatory hematologic markers should be interpreted with consideration of age, sex, and obesity severity.

Keywords: Childhood obesity, inflammatory markers, obesity severity

Öz

Bu bulgular, enflamatuvar hematolojik belirteçlerin yaş, cinsiyet ve obezite şiddeti dikkate alınarak yorumlanması gerektiğini düşündürmektedir.

Anahtar kelimeler: Çocukluk çağı obezitesi, enflamatuvar belirteçler, obezite şiddeti

Introduction

Childhood obesity represents a growing global public health concern and is associated with metabolic, inflammatory, endocrine, and hematologic alterations throughout childhood and adolescence (1). Excess adipose tissue functions not only as an energy storage organ but also as an active endocrine organ that secretes various adipokines and proinflammatory cytokines. Increased adipose tissue accumulation, adipokine imbalance, oxidative stress, and dysregulation of immune system activation contribute to the development of chronic low-grade inflammation in children with obesity (2,3).

Recent studies have demonstrated that obesity-related inflammation may play an important role in the pathogenesis of insulin resistance, metabolic syndrome, cardiovascular disease, and endothelial dysfunction beginning from early life (4,5). Therefore, identifying simple and accessible inflammatory biomarkers in children with obesity has gained increasing clinical importance.

Inflammatory hematological ratios derived from routine complete blood count parameters are inexpensive, practical, and easily accessible markers for evaluating systemic inflammation. Among these, the neutrophil-to-lymphocyte ratio (NLR) and platelet-to-lymphocyte ratio (PLR) have been associated with inflammation and metabolic risk in pediatric obesity; however, findings regarding their relationship with obesity severity remain inconsistent (6-8).

The monocyte-to-lymphocyte ratio (MLR) has been evaluated in a limited number of studies in obese children, and its association with obesity severity is still unclear. Previous studies have suggested that inflammatory hematological ratios may indicate the presence of obesity-related inflammation but may not increase linearly with increasing obesity severity (9-11). In addition, inflammatory responses and hematological parameters may vary according to age and sex (12).

Accordingly, this study aimed to evaluate the relationship between obesity severity and inflammatory hematological ratios, including NLR, PLR, and MLR, in children with obesity while taking age and sex differences into consideration.

Materials and Methods

This single-center cross-sectional observational study was conducted at the Department of Pediatrics of University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital. Pediatric patients aged 5-17 years with obesity who presented to the outpatient clinics between April 1, 2025, and December 1, 2025, were evaluated. All eligible patients who met the predefined inclusion criteria during the study period were included in the study.

Inclusion criteria were age between 5 and 17 years, body mass index-standard deviation score (BMI-SDS) ≥2, and availability of complete blood count data. Exclusion criteria were chronic kidney disease, chronic liver disease, malabsorption syndromes, congenital heart disease, diabetes mellitus, hypothyroidism, Cushing's syndrome, other endocrine disorders, any medication use or supplementation (including vitamin B12, folate, or vitamin D), and the presence of acute or chronic infection.

Ethical Approval

Ethical approval for the study was obtained from the Scientific Research Ethics Committee of University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital on June 25, 2025 (decision no: 2025/231). The extension of the retrospective study period was subsequently approved by the same Ethics Committee at its meeting held on January 14, 2026, under the same study approval (decision no: 2025/231). All procedures were conducted in accordance with the Declaration of Helsinki. The requirement for informed consent was waived due to the retrospective nature of the study.

Anthropometric and Clinical Measurements

Height was measured using a stadiometer with participants wearing light clothing and no shoes. Body weight was measured using a calibrated digital scale. BMI was calculated as weight (kg) divided by height squared (m^2). BMI-SDS values were calculated using the ÇEDD Çözüm/TPEDS Metrics software based on age- and sex-specific reference values for Turkish children reported by Neyzi et al. (13). Obesity was defined as BMI-SDS ≥2 to <3, while severe obesity was defined as BMI-SDS ≥3 according to World Health Organization criteria (14).

Laboratory Assessments

Complete blood count results from venous blood samples collected in the morning after an overnight fast (8-12 hours) were retrospectively retrieved from the hospital information system. Complete blood count parameters were recorded for analysis.

NLR was calculated as the absolute neutrophil count divided by the absolute lymphocyte count. PLR was calculated as the absolute platelet count divided by the absolute lymphocyte count. MLR was calculated as the absolute monocyte count divided by the absolute lymphocyte count. All parameters were derived from complete blood count analysis.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as median (minimum-maximum), and categorical variables as number and percentage [n (%)]. Percentages were calculated within each study group.

Participants were classified into two groups based on BMI-SDS values: Obese (BMI-SDS ≥2 to <3) and severely obese (BMI-SDS ≥3.0). Comparisons were performed between obese and severely obese groups.

Intergroup comparisons were performed using the Mann-Whitney U test for continuous variables and Pearson's chi-square test for categorical variables. To evaluate the association between obesity severity and MLR independently of potential confounding factors, a multivariable linear regression analysis was performed with MLR as the dependent variable and obesity severity group, age, and sex as independent variables.

For age- and sex-stratified NLR analyses, Bonferroni correction was applied to account for multiple comparisons. A p-value <0.05 was considered statistically significant.

Results

A total of 88 individuals with obesity in the pediatric age spectrum (5-17 years) were analyzed. Of these, 65 (73.9%) were classified as obese and 23 (26.1%) as severely obese. The median age of the entire group was 11 (5-17) years, with no significant age difference between groups (p=0.557).

Females constituted a significantly greater proportion of the severely obese group than of the obese group (73.9% vs. 44.6%, p=0.016).

Among hematological parameters, hemoglobin levels were significantly higher in the obese group than in the severely obese group (p=0.026). Basophil counts (p=0.038) and red cell distribution width (p=0.025) also differed significantly between the groups. No statistically significant intergroup differences were observed in neutrophil, lymphocyte, monocyte, platelet, eosinophil, total white blood cell, red blood cell, mean platelet volume, mean corpuscular volume, or plateletcrit values (all p>0.05) (Table 1).

Among inflammatory hematological ratios, MLR values were significantly higher in children with obesity than in those with severe obesity (p=0.003). In contrast, no significant differences were observed in NLR (p=0.176) or PLR (p=0.422). After adjustment for age and sex, greater obesity severity remained independently associated with lower MLR (B=-0.042, standardized β=-0.269, p=0.015) (Table 2).

In the age- and sex-stratified analysis of NLR, an unadjusted intergroup difference was observed among boys aged 5-9 years (p=0.016); however, this finding did not remain statistically significant after Bonferroni correction for multiple comparisons (adjusted p=0.064). No significant differences were observed in girls aged 5-9 years or in boys or girls aged 10-17 years after adjustment for multiple comparisons (Table 3). Sex distribution findings are shown in Figure 1.

Discussion

This study aimed to evaluate the relationship between obesity severity and inflammatory hematological markers in children with obesity, in the context of the well-established association between childhood obesity and low-grade chronic inflammation. Childhood obesity is recognized as a complex disease associated with increased adipose tissue accumulation, adipokine imbalance, oxidative stress, and chronic inflammation (1,2).

Table 1. Demographic and hematological characteristics of the study population

Parameter	Total cases n=88 Median (min-max)	Obese n=65 Median (min-max)	Severely obese n=23 Median (min-max)	p-value
Age, years	11 (5-17)	10.5 (5-16)	13 (5-17)	0.557[1]
Female sex, n (%)	46 (52.3)	29 (44.6)	17 (73.9)	**0.016**[2]
Hemoglobin (g/dL)	12.9 (10.5-16)	13 (10.5-16)	12.5 (10.9-14.3)	**0.026**[1]
Neutrophils (10^3/mm^3)	4.41 (1.47-11.13)	4.44 (1.47-11.13)	4.22 (2.43-8.04)	0.655[1]
Lymphocytes (10^3/mm^3)	2.76 (1.05-6.15)	2.62 (1.05-6.15)	2.9 (1.65-4.26)	0.093[1]
Platelets (10^3/mm^3)	309 (137-525)	309 (137-482)	309 (181-525)	0.525[1]
Monocytes (10^3/mm^3)	0.48 (0.22-1.1)	0.48 (0.22-1.1)	0.44 (0.23-0.65)	0.171[1]
Eosinophils (10^3/mm^3)	0.17 (0-1.05)	0.17 (0-1.05)	0.16 (0.03-0.87)	0.569[1]
Basophils (10^3/mm^3)	0.03 (0.01-0.22)	0.03 (0.01-0.22)	0.03 (0.01-0.06)	**0.038**[1]
Red cell distribution width (%)	13.85 (9-19.7)	13.7 (9-19.7)	14.5 (11.8-17.2)	**0.025**[1]
Mean corpuscular volume (fL)	79.7 (68-94.5)	79.8 (70.6-88)	79 (68-94.5)	0.409[1]
NLR	1.66 (0.46-4.1)	1.7 (0.46-4.1)	1.51 (0.66-3.44)	0.176[1]
PLR	117.33 (49.26-214.43)	119.9 (58.33-214.43)	110.06 (49.26-210.3)	0.422[1]
MLR	0.17 (0.09-0.41)	0.18 (0.11-0.41)	0.15 (0.09-0.36)	**0.003**[1]
Mean platelet volume (fL)	9.15 (7-12.2)	9 (7-12.2)	9.2 (7.3-11.8)	0.868[1]
White blood cells (10^3/mm^3)	7.71 (2.9-16.25)	7.75 (2.9-16.25)	7.5 (4.9-11.2)	0.906[1]
Red blood cells (10^6/mm^3)	4.9 (2.02-6.21)	4.9 (2.02-6.21)	4.89 (3.97-5.53)	0.812[1]
Plateletcrit (%)	0.28 (0.15-0.46)	0.28 (0.15-0.39)	0.26 (0.22-0.46)	0.604[1]

[1]: Mann-Whitney U test, [2]: Pearson's chi-square test, NLR: Neutrophil-to-lymphocyte ratio, PLR: Platelet-to-lymphocyte ratio, MLR: Monocyte-to-lymphocyte ratio

Table 2. Multivariable linear regression analysis of factors associated with MLR

Variable	B	Standardized β	p-value
Sex	-0.005	-0.035	0.756
Obesity severity group	**-0.042**	**-0.269**	**0.015**
Age, years	0.002	0.113	0.305

Dependent variable: MLR. Obesity severity group was coded as 1= obese and 2= morbidly obese. Sex was coded as 1= male and 2= female. B: Unstandardized regression coefficient; MLR: Monocyte-to-lymphocyte ratio

Table 3. Comparison of NLR values according to age group, sex, and obesity severity

Age group	Sex	Obese n	Obese Median (min-max)	Severely obese n	Severely obese Median (min–max)	p-value	Bonferroni-adjusted p
5-9 years	Male	**18**	1.53 (0.70-3.50)	**3**	0.91 (0.84-0.94)	0.016	0.064
	Female	**9**	1.31 (0.84-3.34)	**5**	1.50 (0.96-1.97)	0.841	1.000
10-17 years	Male	**18**	1.70 (0.98-3.80)	**3**	1.56 (0.66-1.58)	0.159	0.636
	Female	**20**	1.90 (0.46-4.10)	**12**	1.90 (1.12-3.44)	0.533	1.000

Bonferroni adjustment was applied for four subgroup comparisons. Adjusted p-values were calculated as p × 4 and capped at 1.000, NLR: Neutrophil-tolymphocyte ratio

Recent studies suggest that obesity-related inflammation begins early in life and may contribute to insulin resistance, endothelial dysfunction, and cardiovascular risk before clinically overt complications become apparent (2,3).

One notable finding of this study was the significantly greater proportion of females in the severely obese group, suggesting that sex-specific biological and hormonal differences may influence obesity severity during

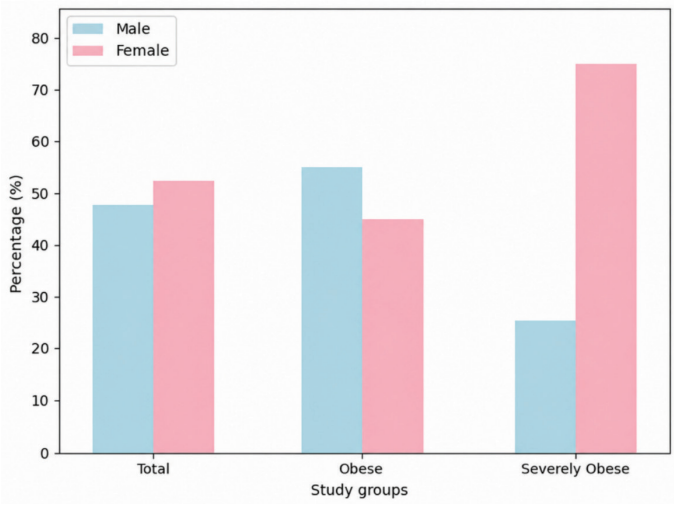

Figure 1. Distribution of sex by obesity severity

childhood and adolescence. Particularly during puberty, increased adipose tissue, elevated leptin levels, and hormonal changes in girls may contribute to more severe obesity (1). This finding is consistent with previous studies emphasizing sex as an important factor influencing obesity severity in pediatric populations (2).

One of the most striking findings of our study was that MLR values were higher in the obese group than in the severely obese group. Importantly, this association remained statistically significant after adjustment for age and sex, indicating that the association between obesity severity and MLR was independent of these potential confounding factors. This finding suggests that inflammation may not increase linearly with increasing obesity severity. While NLR and PLR have been widely investigated in pediatric obesity, data regarding MLR remain limited. Türkkan et al. (6) found no significant difference in MLR between obese adolescents and healthy controls and did not support the utility of inflammatory hematological ratios as biomarkers of adolescent obesity. In contrast, Bala and Bala (8) reported significantly higher MLR values in children with obesity than in normal-weight controls and demonstrated an association between MLR and bone mineral density in the obese group. These differing findings suggest that the relationship between MLR and pediatric obesity may be complex. The mechanisms underlying the lower MLR values observed in severe obesity in our study remain unclear and require further investigation.

Regarding NLR, our findings highlight the importance of considering age and sex when interpreting NLR in relation to obesity severity. Aydın et al. (9) reported that NLR reference ranges in children vary according to age and sex. Likewise, Aydin et al. (3) demonstrated that NLR may be influenced by age- and sex-related differences. In the present study, although an unadjusted difference in NLR was observed among boys aged 5-9 years, this finding did not remain statistically significant after Bonferroni correction for multiple comparisons. Therefore, this subgroup finding should be interpreted cautiously and does not provide sufficient evidence for an independent association between obesity severity and NLR.

The absence of significant differences in PLR, NLR, and most other hematological parameters according to obesity severity suggests that not all inflammatory hematological indices reflect obesity severity to the same extent. Previous studies have shown that although NLR and PLR may indicate increased inflammation in obese children, their ability to distinguish obesity severity remains limited (4,5). Additionally, mean platelet volume and platelet indices may change in obese children, but these changes are not always parallel to obesity severity (10,11).

Another important finding of this study was that hemoglobin levels were higher in the obese group than in the severely obese group. This finding suggests that hematological parameters may not change proportionally with increasing obesity severity. Previous studies have reported associations between obesity and alterations in hematological parameters in children (15,16). In addition, Furuncuoğlu et al. (7) demonstrated that BMI was associated with several complete blood count parameters, although hemoglobin levels did not differ significantly across BMI categories. Therefore, the lower hemoglobin levels observed in the severely obese group in our study may reflect differences in hematological profiles according to obesity severity; however, the underlying mechanisms remain unclear and require further investigation.

Study Limitations

Certain methodological constraints should be acknowledged. The single-center design and modest sample size may restrict the external validity of the results. Second, due to its cross-sectional design, causal relationships between obesity severity and inflammatory hematological parameters cannot be established. Third, inflammatory status was evaluated only through complete blood count-derived ratios, and additional inflammatory biomarkers such as high-sensitivity C-reactive protein or cytokine levels were not assessed. In addition, the absence of a healthy control group limited the ability to compare

inflammatory hematological parameters between obese and non-obese children. Finally, pubertal stage was not evaluated separately, although age- and sex-related differences were considered, which may have influenced the hematological parameters.

Conclusion

Obesity severity was not consistently associated with inflammatory and hematological markers in children with obesity. MLR was lower in children with severe obesity, and this association remained significant after adjustment for age and sex. In contrast, NLR did not differ significantly according to obesity severity after correction for multiple comparisons. These findings suggest that the associations between obesity severity and inflammatory hematological markers may vary across different markers and highlight the importance of considering age, sex, and potential confounding factors in their interpretation. Larger prospective studies are needed to clarify the underlying inflammatory mechanisms.

Ethics

Ethics Committee Approval: Ethical approval for the study was obtained from the Scientific Research Ethics Committee of University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital on June 25, 2025 (decision no: 2025/231).

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

Acknowledgments

During the preparation of this manuscript, ChatGPT (OpenAI) was used as a language assistance tool. The authors reviewed and approved all content and take full responsibility for the final manuscript.

Footnotes

Authorship Contributions

Concept: A.B., Y.C., Ş.G., Design: A.B., Y.C., Ş.G., Data Collection or Processing: A.B., E.N.S., S.G.Ç., F.N.A.Ç., E.B.K.Ç., Analysis or Interpretation: A.B., F.N.A.Ç., Literature Search: A.B., E.N.S., S.G.Ç., E.B.K.Ç., Writing: A.B., E.N.S., S.G.Ç., F.N.A.Ç., E.B.K.Ç., Y.C., Ş.G.

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

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

References

1. Esen İ, Ökdemir D. Childhood obesity: definition, etiology and clinical evaluation. Firat Med J. 2018;23(Suppl):92-99.
2. Derhem B, Karahan İ. Pediyatrik obezite ilişkili erken enflamasyon. Çocuk Enfeksiyon Dergisi. 2023;17(1):9-13.
3. Aydin M, Yilmaz A, Donma MM, Tulubas F, Demirkol M, Erdogan M, et al. Neutrophil/lymphocyte ratio in obese adolescents. North Clin Istanb. 2015;2(2):87-91.
4. Dilek TD, Bostan Gayret Ö, Kılınç S, Erol M, Yiğit Ö, Mete F. The assessment of the neutrophil-lymphocyte ratio and platelet-lymphocyte ratio in dyslipidemic obese children. Bagcilar Med Bull. 2019;4(3):56-60.
5. Yavuz S, Oktay C, Kaçmaz H, Çalışkan AR, Uygun H, Turgut M. Are neutrophil-to-lymphocyte ratio and platelet-to-lymphocyte ratio correlated with hepatosteatosis in obese children? Firat Med J. 2023;28(4):257-261.
6. Türkkan E, Dağ NÇ, Arabacı Ç, Dikker O, Dağ H. Evaluation of inflammatory hematological ratios (NLR, PLR, MLR and monocyte/HDL-cholesterol ratio) in obese adolescents. Iberoam J Med. 2022;4(1):11-17.
7. Furuncuoğlu Y, Tulgar S, Doğan AN, Çakar S, Tulgar YK, Çakıroğlu B. How obesity affects the neutrophil/lymphocyte and platelet/lymphocyte ratio, systemic immune-inflammatory index and platelet indices: a retrospective study. Eur Rev Med Pharmacol Sci. 2016;20:1300-1306.
8. Bala MM, Bala KA. Bone mineral density and complete blood count ratios in children and adolescents with obesity. Eur Rev Med Pharmacol Sci. 2022;26(1):249-256.
9. Aydın İ, Ağıllı M, Aydın FN, Gülcan Kurt Y, Çaycı T, Taş A, et al. The reference ranges of neutrophil-lymphocyte ratio in different age groups. Gulhane Med J. 2015;57:414-418.
10. Emeksiz HC. Mean platelet volume in obese prepubertal children. Akdeniz Med J. 2018;4(1):54-59.
11. Karakurt N, Peltek Kendirci HN. Does body mass index affect mean platelet volume in childhood and adolescence? Turk J Clin Lab. 2019;10:247-250.
12. Delibaş C, Öner C, Çetin H, Kaya Tuna M, Şimşek EE. The relationship of body fat ratio with mean platelet volume, neutrophil/lymphocite and platelet/lymphocyte ratios in obese individuals. Jour Turk Fam Phy. 2022;13(4):155-161.
13. Neyzi O, Bundak R, Gökçay G, Günöz H, Furman A, Darendeliler F, et al. Reference values for weight, height, head circumference, and body mass index in Turkish children. J Clin Res Pediatr Endocrinol. 2015;7(4):280-293.
14. de Onis M, Onyango AW, Borghi E, Siyam A, Nishida C, Siekmann J. Development of a WHO growth reference for school-aged children and adolescents. Bull World Health Organ. 2007;85(9):660-667.
15. Jamshidi L, Seif A. Association between obesity, white blood cell and platelet count. Zahedan J Res Med Sci. 2017;19(2):e4955.
16. Jeong HR, Lee HS, Shim YS, Hwang JS. Positive associations between body mass index and hematological parameters, including RBCs, WBCs, and platelet counts, in Korean children and adolescents. Children (Basel). 2022;9(1):109.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):333-340
DOI: 10.4274/BMB.galenos.2026.85547

Regional Anesthesia on YouTube: Are the Most Viewed Videos the Most Valuable? A Cross-sectional Content Analysis

YouTube'da Rejyonel Anestezi: En Çok İzlenen Videolar En Değerli Olanlar mı? Kesitsel Bir İçerik Analizi

Merve Bulun Yediyıldız[1], Kübra Taşkın[1], İrem Durmuş[1], Hülya Yılmaz Ak[1], Barış Sandal[2], Duygu Demiriz Gülmez[3]

[1]University of Health Sciences Turkey, Kartal Dr. Lütfi Kırdar City Hospital, Department of Anesthesiology and Reanimation, İstanbul, Turkey
[2]İstanbul University-Cerrahpaşa, Cerrahpaşa Faculty of Medicine, Department of Biostatistics, İstanbul, Turkey
[3]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Anesthesiology and Reanimation, İstanbul, Turkey

Abstract

Objective: YouTube is widely used in medical education, particularly for procedural skills such as regional anesthesia. However, concerns remain about the quality and reliability of its content. This study aimed to assess the educational value, reliability, and viewer engagement of regional anesthesia videos on YouTube.

Method: A total of 174 English-language videos were analyzed cross-sectionally. Five anesthesiologists independently evaluated each video using the DISCERN instrument, global quality score (GQS), and JAMA Benchmark Criteria. Viewer interaction was quantified through like ratio, view ratio, and video power index (VPI). Statistical analysis employed the Kruskal-Wallis test and Spearman's correlation.

Results: The median scores were 57 [50-64] for DISCERN, 4 [3-4] for GQS, and 3 [3-3] for JAMA. Video source was significantly associated with DISCERN Q#16, DISCERN, GQS, JAMA, View Ratio, and VPI scores (p=0.010, p=0.005, p=0.048, p<0.001, p<0.001, and p<0.001, respectively), with videos from educational organizations and medical device companies generally showing higher quality scores. Ultrasound-guided videos performed best, while those from anonymous sources and neurostimulator-only techniques had the lowest scores. Correlations between quality and engagement were

Öz

Amaç: YouTube, özellikle rejyonel anestezi gibi girişimsel becerilerin öğrenilmesinde, tıp eğitimi açısından yaygın olarak kullanılan bir platformdur. Bununla birlikte, içeriklerin niteliği ve güvenilirliği konusunda çeşitli kaygılar devam etmektedir. Bu çalışmada, YouTube'daki rejyonel anestezi videolarının eğitsel değeri, güvenilirliği ve izleyici etkileşimi açısından değerlendirilmesi amaçlandı.

Yöntem: İngilizce dilindeki toplam 174 video kesitsel olarak incelendi. Her video, beş anesteziyolog tarafından DISCERN aracı, küresel kalite puanı (GQS) ve JAMA Benchmark Kriterleri kullanılarak bağımsız biçimde değerlendirildi. İzleyici etkileşimi; beğeni oranı, görüntülenme oranı ve video güç indeksi (VPI) ile nicel olarak ölçüldü. İstatistiksel analizde Kruskal-Wallis testi ve Spearman korelasyon analizi kullanıldı.

Bulgular: Ortanca puanlar DISCERN için 57 [50–64], GQS için 4 [3-4] ve JAMA için 3 [3-3] olarak bulundu. Video kaynağı ile DISCERN Q#16, DISCERN, GQS, JAMA, görüntülenme oranı ve VPI skorları arasında anlamlı ilişki saptandı (sırasıyla p=0,010, p=0,005, p=0,048, p<0,001, p<0,001 ve p<0,001); eğitim kuruluşları ve tıbbi cihaz şirketleri tarafından yüklenen videolar genel olarak daha yüksek kalite skorlarına sahipti. Ultrason rehberliğinde hazırlanan videolar

Address for Correspondence: Merve Bulun Yediyıldız, MD, University of Health Sciences Turkey, Kartal Dr. Lütfi Kırdar City Hospital, Department of Anesthesiology and Reanimation, İstanbul, Turkey
E-mail: mervebulun@gmail.com **ORCID:** orcid.org/0000-0002-1809-0791
Received: 11.03.2026 **Accepted:** 06.06.2026 **Epub:** 16.06.2026 **Publication Date:** 29.09.2026

Cite this article as: Bulun Yediyıldız M, Taşkın K, Durmuş İ, Yılmaz Ak H, Sandal B, Demiriz Gülmez D. Regional anesthesia on YouTube: are the most viewed videos the most valuable? A cross-sectional content analysis. Bagcilar Med Bull. 2026;11(3):333-340

Abstract

weak, though view ratio and VPI showed moderate associations with educational value.

Conclusion: The educational quality of YouTube videos on regional anesthesia is inconsistent. Professional, ultrasound-based content offers superior value, while "likes" are unreliable indicators of quality. Evidence-based, peer-reviewed contributions are needed to optimize open-access medical education.

Keywords: Education measurement, medical education, nerve block, ultrasonography, YouTube

Öz

en yüksek performansı gösterirken, anonim kaynaklardan yüklenen videolar ile yalnızca nörostimülatör temelli teknikleri içeren videolar en düşük puanlara sahipti. Kalite ölçütleri ile izleyici etkileşimi arasındaki korelasyonlar genel olarak zayıf bulundu; bununla birlikte, görüntülenme oranı ve VPI'nin eğitsel değer ile orta düzeyde ilişkili olduğu saptandı.

Sonuç: YouTube'daki rejyonel anestezi videolarının eğitsel kalitesi tutarsızdır. Profesyonel nitelikte ve ultrason temelli içerikler daha yüksek eğitsel değer sunarken, beğeni sayıları kaliteyi göstermede güvenilir bir ölçüt değildir. Açık erişimli tıp eğitiminde niteliğin artırılabilmesi için kanıta dayalı ve hakem değerlendirmesinden geçmiş içerik üretimine gereksinim vardır.

Anahtar kelimeler: Eğitim ölçümü, sinir bloğu, tıp eğitimi, ultrasonografi, YouTube

Introduction

Due to the increasing global demand for accessible medical knowledge video sharing platforms, particularly YouTube, have become one of the primary sources of education for healthcare providers and the general public. Although YouTube is considered an educational platform, the quality, accuracy, and reliability of its medical content are still in question (1,2).

Regional anesthesia has become a popular search term as its use in surgical procedures has expanded. It is a technically complex procedure and the guidance techniques are evolving quickly (e.g., ultrasound versus neurostimulation) (3-5). Due to their visual and skill-based nature, video demonstrations have become an integral part of residency training and continuing medical education in this field.

This study evaluates the educational quality, reliability and viewer engagement of YouTube videos related to regional anesthesia using DISCERN, the global quality score (GQS), the JAMA Benchmark Criteria and viewer engagement metrics (6-8). We aimed to determine whether video popularity aligns with educational merit and whether video source or guidance method is associated with higher-quality content.

Materials and Methods

The present cross-sectional descriptive study was conducted to evaluate the quality, reliability and educational value of YouTube videos related to regional anesthesia. The independent review and assessment of all eligible videos was conducted from February 17 to March 27, 2025. The video search was performed using the keyword "regional anesthesia" in the YouTube search bar. In order to prioritize content that has been viewed a significant number of times, the default filter setting, which was previously set to "sort by relevance" has been changed to "sort by view count". All searches were conducted via a web browser after clearing cache, cookies and search history and without logging into any personal account. This method was employed to minimize algorithm-driven personalization of search results.

A preliminary collection of 326 videos exceeding 60 seconds in duration was identified. Exclusion criteria included non-English language, irrelevant content including central neuraxial blocks), duplicate entries, muted or soundless videos, animal-based demonstrations and videos aimed at medical specialties that were not related to anesthesiology (e.g., dentistry, veterinary medicine, ophthalmology or obstetric surgery). Moreover, five videos that had accumulated fewer than 10 views despite being online for over a year were excluded on the grounds of minimal exposure. After applying these criteria, 174 videos remained for final analysis, including 44 (25.3%) upper-extremity peripheral nerve blocks, 52 (29.9%) lower-extremity nerve blocks and 78 (44.8%) truncal/fascial plane blocks and mixed regional anesthesia education videos. The primary analytical framework centered on block videos, which were classified according to the guidance method as ultrasound, neurostimulation, or both. Consequently, central neuraxial blocks were not included in this study.

Each video was independently assessed by five anesthesiologists who were selected on the basis of active clinical practice in regional anesthesia. All reviewers had at least five years of post-residency anesthesia experience and routine clinical experience with ultrasound-guided and/or neurostimulator-assisted upper-extremity, lower-

extremity and truncal/fascial-plane techniques. The reviewers were unaware of the evaluations conducted by their peers and were instructed to evaluate each video independently, without collaboration, discussion or the use of supplementary reference materials. This approach was designed to reduce potential bias and ensure consistency in judgment across evaluators. In order to assess the educational quality and reliability of each video, three validated scoring tools were employed: The DISCERN instrument, the GQS and the JAMA benchmark criteria.

The DISCERN tool, developed by the University of Oxford, is a widely utilized instrument for evaluating the quality of health information with a particular focus on treatment options. The scale used to assess the instrument's reliability, treatment-specific content and overall quality consists of 15 questions, each rated on a scale from 1 to 5. This results in a total score ranging from 15 to 75. The videos were then categorized based on their total score with the following classifications: Excellent (63-75), good (51-62), fair (39-50), poor (27-38) or very poor (15-26) (9,10).

The GQS is a five-point Likert-type scale that provides a structured yet subjective assessment of the educational value, clarity and organization of multimedia health content. This score is indicative of the efficacy with which a video conveys pertinent and comprehensible information to its intended audience (7,11).

The JAMA benchmark criteria, originally proposed by Silberg et al., comprise four binary items evaluating authorship, attribution, disclosure and currency. Each criterion is assigned a score of either 1 (present) or 0 (absent), with a maximum total score of 4 indicating a high degree of transparency and credibility in the content (12).

For each included video, the following metadata were extracted and recorded in Microsoft Excel (Microsoft Corp., Redmond, WA, USA): Video title, total view count, duration (in seconds), upload date, evaluation date (used to calculate time since upload in days), number of likes and dislikes and total number of comments. Furthermore, three viewer engagement metrics were calculated to allow for standardized comparison across videos with varying visibility and age:

Like ratio (%) = (Number of Likes × 100) ÷ (Likes + Dislikes)

View ratio = Total views ÷ Days since upload

Video power index (VPI) = (Like ratio × View ratio) ÷ 100

The VPI, developed by Erdem et al., is a quantitative metric that evaluates a video's popularity and audience engagement on social media. These indices enabled a comparative evaluation of each video's popularity and audience interaction, independent of its publication date (13).

Ethical approval and informed consent were not required for this study because it was based exclusively on publicly available data (YouTube videos) that contained no human participants or identifiable personal information.

Statistical Analysis

The data obtained from the evaluated YouTube videos were recorded and analyzed using IBM SPSS Statistics version 31.0 (IBM Corp., Armonk, NY, USA).

The distribution of continuous variables was assessed using the Shapiro-Wilk test. Descriptive statistics were presented according to the distributional characteristics of the variables. Normally distributed continuous variables were summarized as mean ± standard deviation, whereas non-normally distributed continuous variables were summarized as median [interquartile range (IQR)]. In the group comparison tables, when a variable did not conform to normal distribution in at least one comparison group, the same variable was presented as median [IQR] across all groups to maintain a consistent row-wise presentation. For descriptive completeness, minimum and maximum values were also reported where appropriate. Since several variables deviated from normality, non-parametric statistical methods were used for group comparisons. The Kruskal-Wallis test was used to compare DISCERN Q#16, DISCERN, GQS, JAMA, like ratio, view ratio, and VPI scores across video source categories and guidance method groups. When a significant difference was detected, post hoc pairwise comparisons were performed using the Bonferroni-corrected Mann-Whitney U test. The relationships between educational quality scores and viewer engagement metrics were assessed using Spearman's rank correlation coefficient (ρ). All statistical tests were two-sided, and a p-value of less than 0.05 was considered statistically significant.

Results

A total of 326 videos were initially retrieved from YouTube using the keyword "regional anesthesia", applying a minimum video duration threshold of 60 seconds. Following the screening process, 152 videos were excluded based on predefined criteria (Table 1). The following results summarize the content characteristics, quality assessment scores, and engagement metrics of

the 174 regional anesthesia videos included in the final analysis.

Descriptive analysis revealed marked variability across both video engagement metrics and quality scores. The median number of views was 4102 [649-25659], while the median number of likes and comments were 57 [10-269] and 2 [0-13], respectively. The median DISCERN, GQS, and JAMA scores were 57 [50-64], 4 [3-4], and 3 [3-3], respectively. The median view ratio was 3.27 [0.59-16.61], and the median VPI was 3.68 [0.62-17.23] (Table 2).

When videos were compared according to source, medical device company videos had the highest median DISCERN score (61 [50-75]) and GQS score (5 [3-5]), whereas videos categorized as "others" had the lowest DISCERN and GQS scores (50 [4257] and 3 [2-4], respectively). JAMA scores were also highest in medical device company videos (4 [3-4]). Educational organization videos showed the highest viewer engagement, with a median view ratio of 24.70 [2.41-77.51] and a median VPI of 24.49 [2.28-75.50]. Significant differences were observed across video source groups for DISCERN Q#16, DISCERN, GQS, JAMA, like ratio, view ratio, and VPI scores (p=0.010, p=0.005, p=0.048, p<0.001, p<0.001, p<0.001, and p<0.001, respectively) (Table 3).

According to guidance method, neurostimulation-only videos had the lowest median DISCERN Q#16, DISCERN, and GQS scores (3 [2-4], 52 [42-55], and 3 [2-4], respectively). In comparison, ultrasound-guided videos had higher median DISCERN Q#16, DISCERN, and GQS scores (4 [3-5], 59 [52-66], and 4 [3-5], respectively), while videos using both ultrasound and neurostimulation had a median DISCERN score of 60 [52-64]. Significant differences were observed in DISCERN Q#16, DISCERN, GQS, JAMA, view ratio, and VPI scores (p<0.001, p<0.001, p<0.001, p<0.001, p=0.042, and p=0.016, respectively). Like ratio did not differ significantly among guidance method groups (p=0.411) (Table 4).

DISCERN Q#16 showed strong positive correlations with DISCERN and GQS scores (ρ=0.883 and ρ = 0.947, respectively; both p<0.001), and DISCERN was also strongly correlated with GQS (ρ=0.877, p<0.001). View ratio was positively correlated with DISCERN Q#16, DISCERN, and GQS scores (ρ=0.342, p<0.001; ρ=0.235, p=0.002; and ρ=0.271, p<0.001, respectively). Similarly, VPI was positively correlated with DISCERN Q#16, DISCERN, and GQS

Table 1. Categories and number of excluded videos based on relevance, content type, language and technical criteria during the screening of regional anesthesia videos on YouTube

Category	Number
Educational content for veterinary	1
Educational content for emergency medicine physicians	3
Educational content for algology	16
Educational content for dentists	10
Educational content for ophthalmic surgeons	3
Educational content for obstetric surgeons	1
Educational content for nurses	5
Educational content for orthopedic surgeons	15
Educational content for medical students	3
Patient information videos	33
Non-English language videos	26
Irrelevant to the topic	16
Product promotion videos	15
Low-view videos (<10 views, >1 year online)	5
Total	152

Table 2. Summary of descriptive characteristics and quality assessment scores for included videos

Variable	Minimum	Maximum	Mean	Standard deviation	Median [IQR]
Views	11	352727	29694.71	58270.76	4102 [649-25659]
Likes	0	3900	302.61	578.89	57 [10-269]
Dislikes	0	55	6.03	11.33	0 [0-6]
Comments	0	133	10.91	19.72	2 [0-13]
DISCERN Q#16	1	5	3.68	1.04	4 [3-5]
DISCERN	15	75	56.15	11.60	57 [50-64]
GQS	1	5	3.61	1.02	4 [3-4]
JAMA total	1	4	3.09	0.54	3 [3-3]
Like ratio (%)	85.71	100.00	98.57	2.61	100 [98.13-100]
View ratio	0.017	353.51	23.36	48.12	3.27 [0.59-16.61]
VPI	0.017	347.98	23.57	47.83	3.68 [0.62-17.23]

Data are presented as mean ± standard deviation and median [IQR]. IQR: Interquartile range, VPI: Video power index, GQS: Global quality score

Table 3. Comparison of quality and engagement metrics based on the source of regional anesthesia videos on YouTube (n=174)

	Educational organization on anesthesiology (n=68, 39.1%)	Medical education platform (general-not just anesthesia) (n=25, 14.4%)	Training physician (n=56, 32.2%)	Medical device company (n=11, 6.3%)	Others (n=14, 8%)	p
DISCERN Q#16	4 [3-5]	4 [3-5]	3 [3-4]	5 [3-5]	3 [2-4]	**0.010**
DISCERN	60 [54-65]	54 [51-61]	54 [44-61]	61 [50-75]	50 [42-57]	**0.005**
GQS	4 [3-5]	3 [3-4]	3 [3-4]	5 [3-5]	3 [2-4]	**0.048**
JAMA	3 [3-3]	3 [3-3]	3 [3-3]	4 [3-4]	3 [3-3]	**p<0.001**
Like ratio (%)	98.57 [97.61-100.00]	100 [97.96-100.00]	100 [99.33-100.00]	100.00 [100.00 -100.00]	100.00 [100.00-100.00]	**p<0.001**
View ratio	24.70 [2.41-77.51]	1.40 [0.36-4.15]	2.11 [0.57-7.16]	1.08 [0.51-2.30]	1.13 [0.30-3.03]	**p<0.001**
VPI	24.49 [2.28-75.50]	1.52 [0.38-4.88]	2.32 [0.60-7.07]	1.10 [0.81-2.79]	1.08 [0.30-3.03]	**p<0.001**

Values are presented as median (interquartile range). Comparisons among video source groups were performed using the Kruskal-Wallis test. When a significant difference was detected, post-hoc pairwise comparisons were performed using the Bonferroni-corrected Mann-Whitney U test, VPI: Video power index, GQS: Global quality score

Table 4. Comparison of quality and engagement metrics according to video guidance method (ultrasound-guided, neurostimulator-guided, or both) in regional anesthesia videos (n=174)

	Both (n=41, 23.6%)	Neurostimulation (n=36, 20.7%)	USG (n=97, 55.7%)	p
DISCERN Q#16	4 [3-5]	3 [2-4]	4 [3-5]	**p<0.001**
DISCERN	60 [52-64]	52 [42-55]	59 [52-66]	**p<0.001**
GQS	4 [3-5]	3 [2-4]	4 [3-5]	**p<0.001**
JAMA	3 [3-4]	3 [3-3]	3 [3-3]	**p<0.001**
Like ratio (%)	100.00 [98.48-100.00]	100.00 [98.19-100.00]	99.83 [97.99-100.00]	0.411
View ratio	1.19 [0.29-4.48]	4.25 [1.36-9.31]	5.01 [0.60-49.77]	0.042
VPI	1.11 [0.29-4.48]	4.25 [1.36-9.16]	6.51 [0.68-51.88]	0.016

Values are presented as median (interquartile range). Comparisons among guidance method groups were performed using the Kruskal-Wallis test. When a significant difference was detected, post-hoc pairwise comparisons were performed using the Bonferroni-corrected Mann-Whitney U test, VPI: Video power index, GQS: Global quality score, USG: Ultrasonography

scores (ρ=0.336, p<0.001; ρ=0.224, p=0.003; and ρ=0.262, p=0.001, respectively). In contrast, like ratio showed only a weak negative correlation with DISCERN Q#16 (ρ=*0.178, p=0.021) and was not significantly correlated with DISCERN, GQS, or JAMA scores (Table 5).

Discussion

This study provides a comprehensive evaluation of the educational quality, reliability and viewer engagement of YouTube videos related to regional anesthesia. It uses three respected tools—DISCERN, GQS and the JAMA Benchmark Criteria—along with engagement measurements such as like ratio, view ratio and VPI. The findings reveal the complex relationship between content quality and viewer popularity on open-access platforms and underscore the variability and inconsistency of publicly available educational materials.

The results demonstrate that, although the average video quality ranged from fair to good, the content spectrum was wide. Some videos scored as "excellent," while others fell into the "very poor" category. This type of variance is normal for any unregulated digital platform, such as YouTube, which hosts videos that are neither peer-reviewed nor editorially controlled. Concurrent findings from prior studies of medical education content indicate that such heterogeneity stems from open-access policies and a paucity of curation (1,4). These findings underscore the necessity for content governance and professional involvement in ensuring educational value.

One of the most important findings of our analysis is the significant impact of the video source on quality and reliability metrics. Videos uploaded by academic institutions and medical device companies received higher DISCERN, GQS, and JAMA scores than those published by

Table 5. Spearman correlation coefficients between quality scores and video popularity metrics

	DISCERN Q#16	DISCERN	GQS	JAMA	Like ratio (%)	View ratio	VPI
DISCERN Q#16	–						
DISCERN	**ρ=0.883 p<0.001**	–					
GQS	**ρ=0.947 p<0.001**	**ρ=0.877 p<0.001**	–				
JAMA	**ρ=0.302 p<0.001**	**ρ=0.345 p<0.001**	**ρ=0.324 p<0.001**	–			
Like ratio (%)	**ρ=-0.178 p=0.021**	ρ=-0.104 p=0.179	ρ=-0.131 p=0.089	ρ=-0.044 p=0.569	–		
View ratio	**ρ=0.342 p<0.001**	**ρ=0.235 p=0.002**	**ρ=0.271 p<0.001**	**ρ**=-0.097 p=0.202	**ρ=-0.616 p<0.001**	–	
VPI	**ρ=0.336 p<0.001**	**ρ=0.224 p=0.003**	**ρ=0.262 p=0.001**	ρ=-0.091 p=0.237	**ρ=-0.606 p<0.001**	**ρ=1.000 p<0.001**	–

Correlations were assessed using Spearman's rank correlation analysis. ρ indicates Spearman's correlation coefficient. Two-tailed p-values are presented. Bold values indicate statistically significant correlations (p<0.05), VPI: Video power index, GQS: Global quality score

individual trainers or non-professional sources. Specifically, medical device companies obtained the highest scores across all domains, including DISCERN (61 [50-75]), GQS (5 [3-5]) and JAMA (4 [3-4]). These results indicate that medical device companies adhere to rigorous standards in their content production. Educational organizations demonstrated the highest view ratio (24.70 [2.41-77.51]) and VPI (24.49 [2.28-75.50]), indicative of both quality and dissemination capacity. These results support previous findings that institutional origin is a strong predictor of content quality, suggesting that educational videos associated with established organizations are more likely to adhere to evidence-based practices (13,14).

Post-hoc pair-wise comparisons further corroborated these disparities, with educational organization videos attaining significantly higher DISCERN scores compared to training physician videos (p=0.029). Medical device companies exhibited higher JAMA scores than the other source categories, including educational organizations, medical education platforms, training physicians, and others (p=0.005, p=0.004, p<0.001, and p=0.001, respectively). It is noteworthy that educational institutions exhibited lower like ratios compared to other entities; however, their engagement metrics (VPI, view ratio) were also superior. This prompts the inquiry of whether superficial metrics of approval, such as "likes", may not be the most effective measure of pedagogical value.

The superior performance of ultrasound-guided videos across all assessed dimensions is a particularly noteworthy outcome. The videos obtained high DISCERN Q#16 (4 [3-5]), GQS (4 [3-5]), view ratio (5.01 [0.60-49.77]), and VPI (6.51 [0.68-51.88]) scores, suggesting that they not only provide strong educational value but also resonate well with viewers. This finding aligns with contemporary trends in clinical education practices, where ultrasound has become the prevailing standard in regional anesthesia training. The utilization of ultrasound technology in this context offers distinct advantages, including its capacity to provide clear anatomical visualization and real-time dynamic imaging (4,15). Conversely, neurostimulator-only videos received the lowest quality scores across all measures (DISCERN 52 [42-55]; GQS 3 [2-4]; JAMA =3 [3-3]), reflecting a clear difference in value and relevance of different guidance techniques.

It is noteworthy that the combination of neurostimulation and ultrasound techniques received the highest ratings in the DISCERN and JAMA assessments, with an average score of 60 [52-64] and 3 [3-4], respectively. However, when evaluated against ultrasound-only videos, the combined technique ranked second. This suggests that the combination of techniques may result in the attainment of technical depth and credibility in content, but it may not necessarily lead to its availability or impact on online viewers.

Our study also reveals that viewer engagement metrics, particularly the "like ratio", are not a reliable indicator of the video's educational value. The like ratio exhibited a weak or even negative correlation with DISCERN and GQS scores, while the view ratio and VPI both demonstrated a moderately positive correlation with quality scores. This suggests that the popularity of a video does not necessarily equate to its educational value or the extent to which it is supported by data. These findings are consistent with the conclusions of previous research by Osman et al. (8), which cautioned against using popularity as a proxy for quality in

health-related content. Moreover, the findings of this study corroborate the conclusions of Szmuda et al. (15), who observed that user feedback metrics frequently fall short in accurately reflecting the scientific accuracy or educational value of online videos.

Despite the presence of high-quality videos, our findings indicate that a considerable portion of YouTube content on regional anesthesia remains suboptimal in quality and reliability. Videos from anonymous sources or classified as "others" in particular did poorly on all metrics. These videos may exhibit a lack of authorship transparency, contain information that is outdated, or demonstrate a methodological rigor that is deficient, thereby posing a risk of misinformation. This lack of quality control is alarming and indicates the urgent need for better content governance, especially since YouTube is being used more and more by trainees and students for clinical learning (1,16). In accordance with the findings of Alver et al. (16), it has been determined that YouTube, when utilized as a standalone platform, is inadequate for procedural education due to several limitations, including but not limited to substandard video quality, restricted language options, challenges in accessing contemporary techniques, and an absence of explanatory materials regarding critical anatomical landmarks (17). These issues underscore the imperative for supplementing YouTube content with structured, curriculum-based educational interventions.

Study Limitations

This study has several limitations. First, it included only English-language videos identified using a single keyword ("regional anesthesia"), potentially overlooking relevant content under alternative terms or in other languages. Second, the dynamic nature of YouTube—where views, likes, and recommendations evolve over time—may affect reproducibility. Third, while DISCERN, GQS, and JAMA are validated tools, they may not fully capture the procedural nuances of visual learning. Finally, although all videos were independently assessed by five anesthesiologists with active clinical experience in regional anesthesia, the reviewers' individual procedural volumes and block-specific case numbers were not formally documented. Therefore, some degree of reviewer-related variability in the assessment of block-specific video content cannot be entirely excluded.

Conclusion

This study identifies a significant disparity in the quality and reliability of regional anesthesia videos posted on YouTube. The present study reveals that videos from academic centers and medical device companies, particularly those employing ultrasound guidance, demonstrate consistently higher levels of educational value, credibility, and systematic presentation compared to other sites. However, the use of anonymous or unverified sources can introduce a risk of misinformation, particularly for trainees who rely on the internet for their clinical education.

Specifically, the study corroborates the hypothesis that audience engagement metrics, particularly the Like Ratio, are not an effective indicators of quality in learning. This underscores the necessity for learners and instructors to employ critical appraisal strategies when selecting online learning materials.

Given the centrality of YouTube to medical education, educational institutions, specialty societies, and credible organizations must assume a more proactive role in the production and dissemination of high-quality, evidence-based, open digital learning materials. Subsequent studies must transcend the limitations of quality scoring and assess the tangible value of these videos in terms of learning, procedural competence, and patient outcomes. Cross-platform analyses have the potential to offer further insight into the optimization of digital education across various social media platforms.

Ethics

Ethics Committee Approval: Ethical approval and informed consent were not required for this study because it was based exclusively on publicly available data (YouTube videos) that contained no human participants or identifiable personal information.

Informed Consent: Ethical approval and informed consent were not required for this study because it was based exclusively on publicly available data (YouTube videos) that contained no human participants or identifiable personal information.

Footnotes

Authorship Contributions

Concept: M.B.Y., H.Y.A., Design: M.B.Y., H.Y.A., Data Collection or Processing: M.B.Y., K.T., İ.D., B.S., D.D.G., Analysis or Interpretation: M.B.Y., K.T., İ.D., H.Y.A., B.S., D.D.G., Literature Search: K.T., İ.D., B.S., D.D.G., Writing: M.B.Y., B.S.

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

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

References

1. Nelms MW, Javidan A, Chin KJ, Vignarajah M, Zhou F, Tian C, et al. YouTube as a source of education in perioperative anesthesia for patients and trainees: a systematic review. Can J Anaesth. 2024;71(9):1238-1250. English.
2. Kim JH, Kim HK. Content and quality of YouTube regarding women's health: a scoping review. Korean J Women Health Nurs. 2023;29(3):179-189.
3. Rahman MA, Sahitya IS, Jamal J, Chowdhury TM. Ultrasound-guided regional anesthesia vs landmark technique in resource-limited settings: a comparative study of 50 cases. Clinical Medicine and Health Research Journal (CMHRJ). 2025;6(2):1700-1703.
4. Chequer S. Innovative technologies in regional anaesthesia education: a scoping review. BJA Open. 2025;15:100473.
5. Zhang X, Vanstone RJ, Turbitt L, West S, Harty E. Regional anaesthesia education for consultants and specialists in the UK: a mixed-methods analysis. Br J Anaesth. 2024;132(5):1073-1081.
6. Kasthuri VS, Glueck J, Pham H, Daher M, Balmaceno-Criss M, McDonald CL, et al. Assessing the accuracy and reliability of AI-generated responses to patient questions regarding spine surgery. J Bone Joint Surg Am. 2024;106(12):1136-1142.
7. Yeun AWK.. Evaluation of content quality of online health information by global quality score: a case study of researchers misnaming it and citing secondary sources. Publications. 2025;13(2):23.
8. Osman W, Mohamed F, Elhassan M, Shoufan A. Is YouTube a reliable source of health-related information? A systematic review. BMC Med Educ. 2022;22(1):382.
9. Zhou M, Pan Y, Zhang Y, Song X, Zhou Y. Evaluating AI-generated patient education materials for spinal surgeries: comparative analysis of readability and DISCERN quality across ChatGPT and deepseek models. Int J Med Inform. 2025;198:105871.
10. Cocci A, Pezzoli M, Lo Re M, Russo GI, Asmundo MG, Fode M, et al. Quality of information and appropriateness of ChatGPT outputs for urology patients. Prostate Cancer Prostatic Dis. 2024;27(1):103-108.
11. Zainab A, Sakkour R, Handu K, Mughal S, Menon V, Shabbir D, et al. Measuring the quality of YouTube videos on anxiety: a study using the global quality scale and discern tool. Int J Community Med Public Health. 2023;10(11):4492-4496.
12. Lin JQ, Christopher J, Khela M, Thuo E, Khela H, Chakkalakal D. YouTube videos on lateral epicondylitis often lack high-quality, reliable information. Arthrosc Sports Med Rehabil. 2025;7(3):101150.
13. Tanyıldız B, Oklar M. Evaluating the quality, utility, and reliability of the information in uveitis videos shared on YouTube. Int Ophthalmol. 2023;43(2):549-555.
14. Puthumana JS, Lopez CD, Girard A, Lake IV, Taylor AL, Khoo KH, et al. Evaluating YouTube video quality in orthognathic surgery patient education. Face. 2022;3(1):80-86.
15. Szmuda T, Özdemir C, Fedorow K, Ali S, Soo niewski P. YouTube as a source of information for narcolepsy: a content-quality and optimization analysis. J Sleep Res. 2021;30(2):e13053.
16. Alver S, Ciftci B, Gölboyu BE, Sargolzaeimoghaddam M, Erdoğan C, Alıcı HA, et al. YouTube educational videos for peripheral block learning and application: a survey of Turkish anesthesiologists; how successful. Eurasian J Med. 2023;55(3):208-212.
17. Olgun Y. YouTube as a source of education on piriformis injection: a content, quality, and reliability analysis. BMC Med Educ. 2025;25(1):549.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):341-349
DOI: 10.4274/BMB.galenos.2026.22043

Predictive Performance of Shock Index Derivatives for Mortality in Hospitalized Adult Burn Patients Presenting to the Emergency Department

Acil Servise Başvuran ve Hastaneye Yatırılan Erişkin Yanık Hastalarında Mortalitenin Öngörülmesinde Şok İndeksi Türevlerinin Prediktif Performansı

Muhammed Koru[1], Emin Uysal[2], Burak Demirci[2], Abuzer Coşkun[2], Mücahit Yalçınkaya[3]

[1]Erbaa State Hospital, Clinic of Emergency Medicine, Tokat, Turkey
[2]University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Emergency Medicine, İstanbul, Turkey
[3]Turgutlu State Hospital, Clinic of Emergency Medicine, Manisa, Turkey

Abstract

Objective: Early identification of high-risk burn patients is essential for optimizing clinical management and improving outcomes. Traditional prognostic models mainly rely on burn size and patient age; however, early physiological indicators obtained at emergency department presentation may provide additional prognostic information. This study aimed to evaluate the predictive performance of shock index (SI) derivatives and burn percentage for in-hospital mortality in hospitalized adult burn patients presenting to the emergency department.

Method: This retrospective observational study included adult burn patients who were evaluated in the emergency department and subsequently hospitalized in the burn unit of a tertiary care center. Demographic and clinical data, including burn type, burn depth, total body surface area burned (TBSA), and hospital outcomes, were obtained from electronic medical records. SI, modified shock index (MSI), age shock index (ASI), and burn age shock index (BASI) were calculated using vital signs recorded at initial emergency department presentation. Receiver operating characteristic (ROC) curve analysis was performed to evaluate the predictive performance of these parameters for in-hospital mortality.

Results: A total of 291 hospitalized adult burn patients were included in the study, of whom 33 (11.3%) died during hospitalization. Non-survivors were significantly older and had higher TBSA values than

Öz

Amaç: Yüksek riskli yanık hastalarının erken dönemde belirlenmesi, klinik yönetimin optimize edilmesi ve hasta sonuçlarının iyileştirilmesi açısından büyük önem taşımaktadır. Geleneksel prognostik modeller çoğunlukla yanık yüzdesi ve hasta yaşına dayanmaktadır; ancak acil servise başvuru sırasında elde edilen erken fizyolojik göstergeler ek prognostik bilgi sağlayabilir. Bu çalışmanın amacı, acil servise başvuran ve yanık ünitesine yatırılan erişkin yanık hastalarında şok indeksi (SI) türevleri ve yanık yüzdesinin hastane içi mortaliteyi öngörmedeki performansını değerlendirmektir.

Yöntem: Bu retrospektif gözlemsel çalışmaya acil serviste değerlendirildikten sonra üçüncü basamak bir merkezin yanık ünitesine yatırılan erişkin yanık hastaları dahil edildi. Demografik ve klinik veriler; yanık tipi, yanık derinliği, toplam yanık yüzey alanı (TBSA) ve hastane sonuçları elektronik tıbbi kayıtlardan elde edildi. SI, modifiye şok indeksi (MSI), yaş şok indeksi (ASI) ve yanık yaş şok indeksi (BASI), acil servise ilk başvuru sırasında kaydedilen vital bulgular kullanılarak hesaplandı. Bu parametrelerin hastane içi mortaliteyi öngörme performansını değerlendirmek amacıyla alıcı işletim karakteristiği (ROC) eğrisi analizi yapıldı.

Bulgular: Çalışmaya toplam 291 erişkin yanık hastası dahil edildi ve bunların 33'ü (%11,3) hastanede yaşamını kaybetti. Yaşamını kaybeden hastalar sağ kalanlara göre anlamlı derecede daha yaşlıydı ve TBSA değerleri daha yüksekti (p<0,001). İncelenen tüm

Address for Correspondence: Muhammed Koru, Erbaa State Hospital, Clinic of Emergency Medicine, Tokat, Turkey
E-mail: drmuhammedkoru@gmail.com **ORCID:** orcid.org/0009-0008-0376-6403
Received: 12.03.2026 **Accepted:** 24.07.2026 **Epub:** 05.08.2026 **Publication Date:** 29.09.2026

Cite this article as: Koru M, Uysal E, Demirci B, Coşkun A, Yalçınkaya M. Predictive performance of shock index derivatives for mortality in hospitalized adult burn patients presenting to the emergency department. Bagcilar Med Bull. 2026;11(3):341-349

Abstract

survivors (p<0.001). All evaluated hemodynamic indices were significantly higher among non-survivors (p<0.001). ROC analysis demonstrated that BASI had the highest predictive performance for mortality [area under the curve (AUC): 0.984; 95% confidence interval: 0.972-0.997], followed by TBSA (AUC: 0.939), MSI (AUC: 0.885), SI (AUC: 0.873), and ASI (AUC: 0.849).

Conclusion: SI derivatives and burn percentage were significantly associated with mortality in hospitalized adult burn patients. Among the evaluated parameters, BASI demonstrated the best predictive performance and may represent a practical bedside tool for early mortality risk assessment in burn patients presenting to the emergency department.

Keywords: Burn age shock index, burn injury, emergency department, mortality, shock index

Öz

hemodinamik indeksler non-survivor yaşamını kaybeden hasta grubunda anlamlı derecede daha yüksek bulundu (p<0,001). ROC analizi sonucunda mortaliteyi öngörmede en yüksek performansı BASI gösterdi [eğri altında kalan alan (AUC): 0,984; %95 güven aralığı: 0,972-0,997]. Bunu sırasıyla TBSA (AUC: 0,939), MSI (AUC: 0,885), SI (AUC: 0,873) ve ASI (AUC: 0,849) izledi.

Sonuç: Şok indeksi türevleri ve yanık yüzdesi, hastaneye yatırılan erişkin yanık hastalarında mortalite ile anlamlı olarak ilişkilidir. İncelenen parametreler arasında BASI en yüksek öngörü performansını göstermiş olup, acil servise başvuran yanık hastalarında erken mortalite risk değerlendirmesi için pratik bir yatak başı araç olabilir.

Anahtar kelimeler: Acil servis, mortalite, şok indeksi, yanık, yanık yaş şok indeksi

Introduction

Burn injuries remain a major cause of morbidity and mortality worldwide and represent a significant public health burden. It is estimated that millions of burn injuries occur each year globally, with a substantial proportion resulting in hospitalization and long-term disability. Severe burns may lead to complex physiological disturbances including systemic inflammatory responses, extensive fluid shifts, and circulatory instability, which may ultimately progress to organ dysfunction and death. Early identification of patients at increased risk of adverse outcomes is therefore crucial for guiding treatment strategies and optimizing clinical management in hospitalized burn patients (1,2).

Several clinical factors have traditionally been associated with mortality in burn patients. Among these, total body surface area (TBSA) burned and patient age are recognized as the most important determinants of outcome. Prognostic scoring systems such as the Baux score and the abbreviated burn severity index (ABSI) integrate these parameters and have been widely used to estimate mortality risk in burn populations (3,4). These scoring systems provide valuable prognostic information; however, they mainly rely on demographic and injury characteristics and may not fully reflect the early physiological status of patients presenting to the emergency department. Early hemodynamic alterations occurring during the initial phase of burn injury may therefore provide additional prognostic information that is not captured by conventional scoring models (5).

The shock index (SI), defined as the ratio of heart rate to systolic blood pressure (SBP), is a simple bedside indicator that reflects hemodynamic compromise and circulatory instability. Initially introduced in emergency medicine, SI has been widely investigated as a prognostic marker in trauma, sepsis, and critically ill patients (6,7). In recent years, several derivatives of the SI have been proposed in order to improve its predictive accuracy. These include the modified shock index (MSI), age shock index (ASI), and burn age shock index (BASI), which incorporate additional physiological or demographic variables to enhance risk stratification (8-10). Although these indices have demonstrated prognostic value in various clinical settings, evidence regarding their predictive performance in burn patients remains limited.

In particular, few studies have evaluated the comparative performance of SI derivatives together with burn severity parameters in cohorts of hospitalized burn patients initially assessed in the emergency department. Understanding the prognostic value of these readily available indices may facilitate early risk stratification and clinical decision-making in burn care. Therefore, the present study aimed to evaluate the predictive performance of SI derivatives and burn percentage for in-hospital mortality in hospitalized adult burn patients presenting to the emergency department.

Materials and Methods

Study Design and Setting

This study was designed as a retrospective observational study conducted at a tertiary care hospital providing specialized treatment for burn injuries. The hospital functions as a regional referral center for burn care and receives patients with varying degrees of burn severity from a broad geographic area. The study population consisted of adult burn patients who were initially evaluated in the emergency department and subsequently hospitalized

in the burn unit. Medical records of eligible patients were retrospectively reviewed using the hospital's electronic medical record system in order to collect demographic, clinical, and outcome-related information.

Study Population

Adult patients aged 18 years and older who were admitted to the burn unit following evaluation in the emergency department between January 1, 2022, and November 30, 2024 were considered eligible for inclusion. Patients were excluded if they were younger than 18 years of age, if they were discharged directly from the emergency department without hospitalization, if their medical records were incomplete, or if the vital sign measurements required for calculating SI parameters were missing. After applying these criteria, a total of 291 hospitalized adult burn patients were included in the final analysis.

Data Collection

Demographic and clinical data were obtained from electronic hospital records and patient charts. The collected variables included age, sex, burn type, burn depth, burn percentage, presence of inhalation injury, presence of additional trauma, admission location, hospital length of stay, and in-hospital mortality. Burn types were classified as thermal, electrical, chemical, or cold injury according to the mechanism of injury documented in the medical records. Burn depth was categorized based on standard clinical burn classifications and included superficial second-degree burns, deep second-degree burns, third-degree burns, and fourth-degree burns.

Burn severity was quantified using TBSA, which was determined according to standard burn assessment methods documented in the patient charts. The presence of inhalation injury and additional traumatic injuries accompanying the burn event were also recorded. The hospitalization location was documented as either ward admission or intensive care unit admission.

Vital signs obtained at the time of the initial emergency department evaluation were used for the calculation of hemodynamic indices. These measurements included heart rate, SBP, and diastolic blood pressure (DBP), which were recorded during the first clinical assessment in the emergency department.

Calculation of SI Derivatives

Several hemodynamic indices were calculated using the recorded vital signs and demographic variables in order to evaluate their prognostic value in hospitalized burn patients. The SI was calculated as the ratio of heart rate to SBP (SI = heart rate/SBP). The MSI was calculated as the ratio of heart rate to mean arterial (MAP) (MSI = heart rate/MAP), where mean arterial pressure was calculated using the standard physiological formula MAP = (SBP + 2 × DBP) / 3. The ASI was obtained by multiplying the SI by the patient's age (ASI = age × SI). The BASI incorporated burn severity into this calculation by multiplying the patient's age, burn percentage expressed as TBSA, and the SI value (BASI = age × TBSA × SI). All indices were calculated using the initial vital signs recorded at the time of emergency department presentation.

Outcome Measures

The primary outcome of the study was in-hospital mortality. Patients were classified as survivors or non-survivors based on their hospital outcomes. Additional analyses evaluated the relationship between SI derivatives and burn characteristics, including burn depth and burn percentage, in order to assess the clinical relevance of these indices in different burn severity groups.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics software (IBM Corp., Armonk, NY, USA). Continuous variables were expressed as mean values with standard deviations, whereas categorical variables were presented as frequencies and percentages. Baseline demographic and clinical characteristics of the study population were summarized descriptively. Comparisons between survivors and non-survivors were performed using the independent samples Student's t-test for continuous variables. The relationship between SI derivatives and burn depth groups was analyzed using One-Way Analysis of Variance (ANOVA) for normally distributed variables and the Kruskal-Wallis test for non-normally distributed variables.

Receiver operating characteristic (ROC) curve analysis was performed to evaluate the predictive performance of SI derivatives and burn percentage for in-hospital mortality. The area under the curve (AUC) was calculated for each parameter, and optimal cut-off values were determined using the Youden index. Sensitivity and specificity values corresponding to these cut-off points were calculated in order to assess the diagnostic performance of each parameter. A p-value of less than 0.05 was considered statistically significant for all analyses.

Ethics Approval

The study protocol was approved by the Institutional Ethics Committee of University of Health Sciences Turkey,

İstanbul Bağcılar Training and Research Hospital (decision no: 2025/01/93/003, date: 17.01.2025), and the study was conducted in accordance with the principles of the Declaration of Helsinki. Because of the retrospective nature of the study and the use of anonymized patient data, the requirement for informed consent was waived by the ethics committee.

Results

Study Population and Baseline Characteristics

A total of 291 adult burn patients who were hospitalized after evaluation in the emergency department were included in the study. Of these patients, 215 (73.9%) were male and 76 (26.1%) were female. Thermal burns constituted the most common type of injury (85.9%), followed by electrical burns (7.2%), chemical burns (6.2%), and cold injuries (0.7%). Regarding burn depth, the majority of patients had deep second-degree burns (70.1%), while 24.4% had third-degree burns and 5.2% had superficial second-degree burns. Only one patient (0.3%) had a fourth-degree burn. Inhalation injury was identified in 13 patients (4.8%), and additional trauma was present in 15 patients (5.2%). In terms of hospitalization location, 155 patients (53.3%) required admission to the intensive care unit (ICU), whereas 136 patients (46.7%) were admitted to the ward (Table 1).

Table 1. Baseline characteristics of hospitalized adult burn patients

Category	Variable	n (%)
Demographics	Total patients	291 (100)
	Male	215 (73.9)
	Female	76 (26.1)
Burn type	Thermal burn	250 (85.9)
	Chemical burn	18 (6.2)
	Electrical burn	21 (7.2)
	Cold injury	2 (0.7)
Burn depth	2nd degree superficial	15 (5.2)
	2nd degree deep	204 (70.1)
	3rd degree	71 (24.4)
	4th degree	1 (0.3)
Clinical features	Additional trauma present	15 (5.2)
	Inhalation injury present	13 (4.5)
Admission location	Ward admission	136 (46.7)
	ICU admission	155 (53.3)

Data are presented as n (%). Percentages were calculated based on the total study population (n=291). ICU: Intensive care unit

Comparison According to in-hospital Mortality

Among the study population, 33 patients died during hospitalization, while 258 patients survived. Non-survivors were significantly older than survivors (53.21±18.36 vs. 39.00±14.99 years, p<0.001). However, length of hospital stay did not differ significantly between the two groups (16.30±13.35 vs. 13.39±16.62 days, p=0.335). All evaluated hemodynamic indices were significantly higher in non-survivors compared with survivors. The SI was significantly elevated in non-survivors (0.921±0.162) compared with survivors (0.711±0.147, p<0.001). Similarly, the MSI was higher in non-survivors (1.302±0.233) than in survivors (0.985±0.210, p<0.001). The ASI also demonstrated a marked difference between the groups (48.55±17.34 vs. 27.68±12.25, p<0.001). The largest difference was observed for the BASI, which was substantially higher among non-survivors (2785.95±1179.24) compared with survivors (568.99±379.14, p<0.001). Additionally, the burn percentage (TBSA) was significantly greater in non-survivors (60.61±20.87%) than in survivors (20.89±11.10%, p<0.001) (Table 2).

SI Derivatives According to Burn Depth

When patients were stratified according to burn depth, significant differences were observed in SI derivatives

Table 2. Clinical variables according to in-hospital mortality

Variable	Non-survivors (n=33) Mean ± SD	Survivors (n=258) Mean ± SD	p-value
Age (years)	53.21±18.36	39.00±14.99	<0.001
Length of hospital stay (days)	16.30±13.35	13.39±16.62	0.335
Shock index	0.921±0.162	0.711±0.147	<0.001
Modified shock index	1.302±0.233	0.985±0.210	<0.001
Age shock index	48.55±17.34	27.68±12.25	<0.001
Burn age shock index	2785.95±1179.24	568.99±379.14	<0.001
Burn percentage (TBSA, %)	60.61±20.87	20.89±11.10	<0.001

Values are presented as mean ± standard deviation (SD). Continuous variables were compared between non-survivors and survivors using the independent samples Student's t-test. Statistical significance was defined as p<0.05, TBSA: Total body surface area burned

and burn percentage. The SI increased with burn severity, from 0.695±0.090 in superficial second-degree burns to 0.789±0.181 in third-degree burns (p=0.008). A similar trend was observed for the MSI, which increased from 0.959±0.132 to 1.102±0.259 across burn depth categories (p=0.005). The ASI and BASI also demonstrated significant increases with increasing burn depth (p=0.002 and p<0.001, respectively). Likewise, the burn percentage (TBSA) was significantly higher in patients with deeper burns (p<0.001) (Table 3). Interpretation of findings related to the fourth-degree burn subgroup should be made cautiously, as only one patient was included in this category. Therefore, measures of variability could not be reliably calculated and statistical comparisons involving this subgroup may not be stable.

ROC Analysis for Prediction of Mortality

ROC curve analysis was performed to evaluate the predictive performance of SI derivatives and burn percentage for in-hospital mortality. Among the evaluated parameters, the BASI demonstrated the highest predictive performance with an AUC of 0.984 [95% confidence interval (CI): 0.972-0.997]. The optimal cut-off value for BASI was 1510.93, yielding a sensitivity of 90.9% and specificity of 96.9%. The burn percentage (TBSA) also showed excellent predictive ability with an AUC of 0.939 (95% CI: 0.889-0.988) and an optimal cut-off value of 37.5%, corresponding to 87.9% sensitivity and 90.7% specificity. Other indices also demonstrated good predictive performance. The MSI had an AUC of 0.885, followed by the SI with an AUC of 0.873, and the ASI with an AUC of 0.849. All ROC analyses were statistically significant (p<0.001) (Table 4, Figure 1).

Discussion

The present study provides three main findings. First, all SI derivatives were significantly associated with in-hospital mortality among hospitalized adult burn patients. Second, burn percentage demonstrated strong predictive performance for mortality. Third, and most importantly, BASI showed the highest discriminative ability among all evaluated parameters. These findings suggest that the integration of age, burn extent, and early hemodynamic status may provide valuable prognostic information during the initial assessment of burn patients in the emergency department.

The prognostic importance of burn size and age has long been recognized in burn care. Classical mortality models such as the Baux score and the ABSI are largely based on these variables and remain widely used in clinical practice (3,4).

Table 3. Shock index derivatives and burn percentage according to burn depth

Variable	2nd degree superficial Mean ± SD	2nd degree deep Mean ± SD	3rd degree Mean ± SD	4th degree (n=1)	p-value
Shock index	0.695±0.090	0.719±0.157	0.789±0.181	0.792	0.008
Modified shock index	0.959±0.132	0.997±0.227	1.102±0.259	1.018	0.005
Age shock index	24.462±8.378	28.503±13.666	35.726±16.378	24.552	0.002
Burn age shock index	294.337±122.210	639.807±547.345	1455.064±1356.119	491.040	<0.001
Burn percentage (TBSA, %)	12.67±5.63	21.97±11.61	38.01±26.12	20	<0.001

Values are presented as mean ± standard deviation (SD), except for the 4th-degree burn group (n=1) which is presented as a single observation value. Comparisons between burn depth groups were performed using One-Way Analysis of Variance for normally distributed variables or the Kruskal-Wallis test for non-normally distributed variables. A p-value <0.05 was considered statistically significant. Standard deviation could not be calculated for the 4th-degree burn group due to the presence of a single patient. Because only one patient had a fourth-degree burn, statistical estimates for this subgroup should be interpreted cautiously and were not considered reliable for subgroup inference, TBSA: Total body surface area burned

Table 4. ROC analysis for predicting in-hospital mortality

Parameter	AUC	95% CI	Cut-off	Sensitivity (%)	Specificity (%)	p-value
Shock index	0.873	0.821-0.924	0.796	90.9	76.4	<0.001
Modified shock index	0.885	0.834-0.936	1.076	93.9	79.5	<0.001
Age shock index	0.849	0.774-0.925	35.53	84.8	76.7	<0.001
Burn age shock index	0.984	0.972-0.997	1510.93	90.9	96.9	<0.001
Burn percentage (TBSA, %)	0.939	0.889-0.988	37.50	87.9	90.7	<0.001

ROC: Receiver operating characteristic, AUC: Area under the curve, TBSA: Total body surface area burned, CI: Confidence interval. Optimal cut-off values were determined using the Youden index. Sensitivity and specificity correspond to the values obtained at the optimal cut-off point. A p-value <0.05 was considered statistically significant

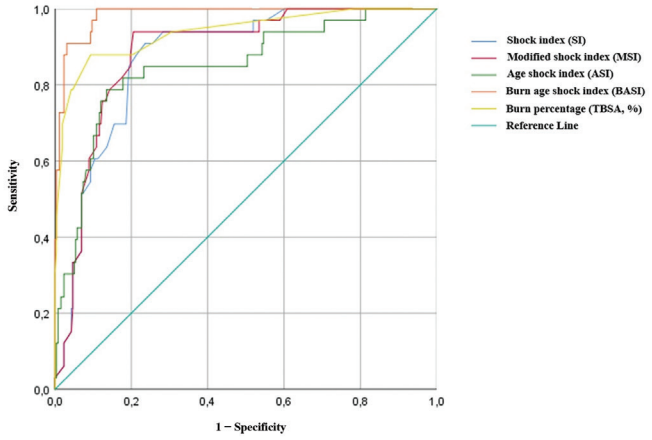

Figure 1. Receiver operating characteristic curves demonstrating the predictive performance of shock index derivatives (SI, MSI, ASI, BASI) and burn percentage for mortality in hospitalized adult burn patients presenting to the emergency department

TBSA: Total body surface area burned

In our cohort, non-survivors were significantly older than survivors and had substantially larger burned body surface areas, which is consistent with the established literature on burn mortality. The strong predictive ability of burn percentage observed in our study also supports the continued clinical relevance of TBSA as one of the most important determinants of outcome in burn patients. Similar findings have been reported in previous burn studies demonstrating that TBSA is strongly associated with mortality risk (11).

Severe burn injury is associated with extensive fluid shifts, systemic inflammatory activation, and circulatory instability, particularly in patients with larger burns (2). For this reason, physiological indices derived from routine vital signs may capture early deterioration that is not fully reflected by anatomical burn characteristics alone. The SI has been widely studied in emergency medicine as a simple marker of hemodynamic compromise. In trauma populations, elevated SI values have been associated with hypovolemic shock, transfusion requirement, and increased mortality risk. Mutschler et al. (12) demonstrated that SI can serve as a rapid indicator of transfusion requirement in injured patients. In addition, a systematic review and meta-analysis by Vang et al. (13) showed that higher SI values were significantly associated with mortality in trauma populations. Dai et al. (14) also reported that different SI derivatives have significant predictive value for early mortality in trauma patients.

In our study, both SI and MSI were significantly higher in non-survivors, indicating that early circulatory stress is closely related to adverse outcomes in burn patients. MSI may be particularly useful because it incorporates mean arterial pressure and therefore may reflect tissue perfusion more comprehensively than SBP alone. Although the literature specifically evaluating MSI in burn cohorts remains limited, its performance in our study suggests that it may represent a clinically relevant marker in this population.

Among all evaluated parameters, BASI demonstrated the highest predictive performance. BASI integrates three major determinants of burn prognosis: Patient age, burn extent, and early hemodynamic status. From a clinical perspective, this combined structure may provide a more comprehensive assessment of physiological stress than isolated variables. Older patients typically have lower physiological reserve, extensive burns impose greater metabolic and inflammatory burden, and abnormal vital signs reflect early circulatory compromise. When these components are combined into a single index, prognostic accuracy may improve.

From a practical bedside perspective, BASI may provide clinicians with a rapid and easily applicable adjunct for early risk stratification at the time of emergency department presentation, as it relies solely on routinely available clinical variables without requiring additional laboratory testing. In contrast to established prognostic models such as the Baux score and the ABSI, which are primarily based on demographic characteristics and burn severity, BASI additionally incorporates early physiological status through the SI and may therefore better reflect acute circulatory stress during the initial phase of burn injury. However, BASI should be considered complementary to established prognostic tools rather than a replacement and should always be interpreted within the broader clinical context and clinical judgment (4,6,15).

Recent studies evaluating SI derivatives in burn patients have reported similar findings. Içer et al. (15) demonstrated that BASI had the highest predictive performance among several shock index-based parameters for mortality prediction in burn patients. In another study conducted by the same research group, the prognostic value of indices such as ASI was shown to vary according to the age distribution of the study population, particularly when pediatric patients were included (16). Since our cohort consisted exclusively of hospitalized adult patients, the

strong predictive performance of BASI observed in our study is consistent with these previous findings.

Male predominance among burn patients has been reported in several national and international cohorts. Regarding national data: Özçetin et al. (17) reported that 59% of burn patients in their cohort were male. Similarly, Şakrak et al. (18) found that 60.2% of hospitalized burn patients were male. In other national studies, Günay et al. (19) reported a male predominance of approximately 70%, and Çınal and Barın (20) observed that 59.7% of burn cases were male. Similar findings have also been reported internationally; Kobayashi et al. (21) documented a male predominance in Japan, while Akerlund et al. (22) reported comparable results in Sweden. The epidemiological characteristics of our cohort were largely compatible with these previously published burn studies. In our study, the proportion of male patients was slightly higher than in many previous reports, which may be related to the metropolitan and industrial environment of our center, where occupational exposure to burn injuries may be more common among men.

Thermal burns were the most common mechanism of injury in our cohort, which is consistent with the epidemiological pattern described in the literature. Previous studies have similarly reported that thermal burns constitute the majority of burn injuries (19,20). The mean burn percentage in our study was also higher than that reported in many general burn series. Previous epidemiological studies have reported average TBSA values ranging approximately between 8% and 14% (23-25). The higher mean TBSA observed in our cohort is most likely explained by the fact that our study included only hospitalized burn patients rather than minor cases treated in outpatient settings.

Another important observation was that SI derivatives and burn percentage increased significantly with burn depth. Deeper burns are generally associated with more extensive tissue injury, greater inflammatory burden, and more severe physiological stress. Previous studies have similarly reported that deeper burns are associated with more complicated clinical courses and prolonged treatment periods (26,27). These findings suggest that burn depth may also be closely related to the physiological instability reflected by shock index-based parameters.

The clinical implications of these findings are noteworthy. In the emergency department, early risk assessment is crucial for determining the level of monitoring, the need for burn unit admission, and the intensity of resuscitative care. SI derivatives can be calculated rapidly from routinely obtained vital signs and basic clinical information. In this context, BASI may serve as a practical adjunct for early bedside risk assessment in hospitalized adult burn patients. However, these indices should be considered complementary to clinical evaluation and established burn severity models rather than replacements for them.

Study Limitations

This study has several limitations. Its retrospective single-center design may limit generalizability and introduces the possibility of documentation bias. Some patients were excluded because of incomplete records, which may have influenced cohort composition. In addition, causes of death during hospitalization may not always have been clearly distinguishable from burn-related complications in retrospective records. Despite these limitations, the study also has important strengths, including the evaluation of multiple SI derivatives in hospitalized adult burn patients and the direct comparison of these indices with burn percentage. Nevertheless, the present study provides valuable data by focusing specifically on hospitalized adult burn patients and by directly comparing multiple SI derivatives with burn percentage in the same cohort. An additional limitation should be acknowledged regarding burn-depth subgroup analyses. Only one patient in the study population had a fourth-degree burn, which limited the statistical robustness of comparisons across burn depth categories. Although this patient was retained to preserve cohort completeness, results related to this subgroup should be interpreted cautiously.

Conclusion

In conclusion, SI derivatives and burn percentage were all associated with mortality in hospitalized adult burn patients presenting to the emergency department, and BASI demonstrated the best predictive performance among the evaluated parameters. These findings suggest that BASI may represent a useful and easily applicable tool for early mortality risk assessment in adult burn patients. Further large-scale prospective multicenter studies are required to validate these findings and to determine the role of SI derivatives in routine burn prognostication. From a clinical perspective, the use of simple bedside indices such as BASI may help emergency physicians identify high-risk burn patients earlier and prioritize intensive monitoring or aggressive resuscitation strategies when necessary. However, external validation studies are needed before routine implementation of BASI in clinical practice.

Ethics

Ethics Committee Approval: The study protocol was approved by the Institutional Ethics Committee of University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital (decision no: 2025/01/93/003, date: 17.01.2025).

Informed Consent: Because of the retrospective nature of the study and the use of anonymized patient data, the requirement for informed consent was waived by the ethics committee.

Acknowledgments

During the preparation of this manuscript, the authors used an AI-assisted language editing tool to improve the clarity, grammar, and readability of the text. After using this tool, the authors carefully reviewed and edited the manuscript and take full responsibility for the content of the published article.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.K., E.U., B.D., A.C., M.Y., Concept: M.K., E.U., B.D., A.C., M.Y., Design: M.K., E.U., B.D., A.C., M.Y., Data Collection or Processing: M.K., E.U., B.D., A.C., M.Y., Analysis or Interpretation: M.K., E.U., B.D., A.C., M.Y., Literature Search: M.K., E.U., B.D., A.C., M.Y., Writing: M.K., E.U., B.D., A.C., M.Y.

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

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

References

1. Peck MD. Epidemiology of burns throughout the world. Part I: distribution and risk factors. Burns. 2011;37(7):1087-1100.
2. Jeschke MG, van Baar ME, Choudhry MA, Chung KK, Gibran NS, Logsetty S. Burn injury. Nat Rev Dis Primers. 2020;6(1):11.
3. Osler T, Glance LG, Hosmer DW. Simplified estimates of the probability of death after burn injuries: extending and updating the baux score. J Trauma. 2010;68(3):690-697.
4. Tobiasen J, Hiebert JM, Edlich RF. The abbreviated burn severity index. Ann Emerg Med. 1982;11(5):260-262.
5. Brusselaers N, Monstrey S, Vogelaers D, Blot S. Outcome prediction in severe burn injury: clinical versus laboratory markers. Eur J Trauma Emerg Surg. 2011;37(5):533.
6. Rady MY, Smithline HA, Blake H, Nowak R, Rivers E. A comparison of the shock index and conventional vital signs to identify acute, critical illness in the emergency department. Ann Emerg Med. 1994;24(4):685-690.
7. Koch E, Lovett S, Nghiem T, Riggs RA, Rech MA. Shock index in the emergency department: utility and limitations. Open Access Emerg Med. 2019;11:179-199.
8. Liu YC, Liu JH, Fang ZA, Shan GL, Xu J, Qi ZW, et al. Modified shock index and mortality rate of emergency patients. World J Emerg Med. 2012;3(2):114-117.
9. Kim SY, Hong KJ, Shin SD, Ro YS, Ahn KO, Kim YJ, et al. Validation of the shock index, modified shock index, and age shock index for predicting mortality of geriatric trauma patients in emergency departments. J Korean Med Sci. 2016;31(12):2026-2032.
10. Olaitan PB, Jiburum BC. Analysis of burn mortality in a burns centre. Ann Burns Fire Disasters. 2006;19(2):59-62.
11. Brusselaers N, Monstrey S, Colpaert K, Decruyenaere J, Blot SI, Hoste EA. Outcome of acute kidney injury in severe burns: a systematic review and meta-analysis. Intensive Care Med. 2010;36(6):915-925.
12. Mutschler M, Nienaber U, Münzberg M, Wölfl C, Schoechl H, Paffrath T, et al.; TraumaRegister DGU. The shock index revisited - a fast guide to transfusion requirement? A retrospective analysis on 21,853 patients derived from the TraumaRegister DGU. Crit Care. 2013;17(4):R172.
13. Vang M, Østberg M, Steinmetz J, Rasmussen LS. Shock index as a predictor for mortality in trauma patients: a systematic review and meta-analysis. Eur J Trauma Emerg Surg. 2022;48(4):2559-2566.
14. Dai G, Lu X, Xu F, Xu D, Li P, Chen X, et al. Early mortality risk in acute trauma patients: predictive value of injury severity score, trauma index, and different types of shock indices. J Clin Med. 2022;11(23):7219.
15. Içer M, Gündüz E, Akkoç MF, Polat D, Özkan H, Bayrak T, et al. The effectiveness of shock indices on prognosis in burn patients admitted to the emergency department. Ulus Travma Acil Cerrahi Derg. 2023;29:786-791.
16. Içer M, Gündüz E, Güloğlu C, Erbatur S, Polat D, Özkan H, et al. Vital indicators in predicting burn mortality: a comparison of shock indices and burn shock indices. Turk J Med Sci. 2023;53(6):1877-1885.
17. Özçetin B, Tihan D, Demirci H, Altıntaş MM, Arayıcı V, Taha A. Two and a half years experience at a new burn center. Turk J Surg. 2012;28(3):146-148.
18. Şakrak T, Köse AA, Karabağlı Y, Çetin C. Our 10-year screening results of patients receiving inpatient treatment in our burn unit. Turk Plast Surg. 2010;18(3):111-115.
19. Günay K, Taviloğlu K, Eskioğlu E, Ertekin C. A study of epidemiology and mortality in burn patients. Ulusal Travma Acil Cerrahi Derg. 1995;1(2):205-208.
20. Çınal H, Barın EZ. Five years of experience in a burn care unit: analysis of burn injuries in 667 patients. Van Medical Journal. 2020;27(1):56-62.
21. Kobayashi K, Ikeda H, Higuchi R, Nozaki M, Yamamoto Y, Urabe M, et al. Epidemiological and outcome characteristics of major burns in Tokyo. Burns. 2005;31(Suppl 1):S3-S11.
22. Akerlund E, Huss FR, Sjöberg F. Burns in Sweden: an analysis of 24,538 cases during the period 1987-2004. Burns. 2007;33(1):31-36.
23. Çiftçi İ, Arslan K, Altunbaş Z, Kara F, Yılmaz H. Epidemiologic evaluation of patients with major burns and recommendations for burn prevention. Ulus Travma Acil Cerrahi Derg. 2012;18(2):105-110.

24. Mzezewa S, Jonsson K, Aberg M, Salemark L. A prospective study on the epidemiology of burns in patients admitted to the Harare burn units. Burns. 1999;25(6):499-504.

25. Tung KY, Chen ML, Wang HJ, Chen GS, Peck M, Yang J, et al. A seven-year epidemiology study of 12,381 admitted burn patients in Taiwan. Burns. 2005;31(Suppl 1):12-17.

26. Zheng Y, Lin G, Zhan R, Qian W, Yan T, Sun L, et al. Epidemiological analysis of 9,779 burn patients in China: an eight-year retrospective study at a major burn center. Exp Ther Med. 2019;17(4):2847-2854.

27. Anlatici R, Ozerdem OR, Dalay C, Kesiktaş E, Acartürk S, Seydaoglu G. A retrospective analysis of 1083 Turkish patients with serious burns. Burns. 2002;28(3):231-237.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):350-358
DOI: 10.4274/BMB.galenos.2026.27147

Factors Associated with the Recurrence of Hemoptysis Following Bronchial Artery Embolization

Bronşiyal Arter Embolizasyonu Sonrasında Nüks Hemoptizi ile İlişkili Faktörler

Eda Cingöz[1], Ali Dablan[2], Mustafa Fatih Arslan[2], Oğuzhan Türksayar[2], Mehmet Cingöz[2], Çağrı Erdim[2], İlhan Nahit Mutlu[2], Tevfik Güzelbey[3], Özgür Kılıçkesmez[4]

[1]University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Radiology, İstanbul, Turkey
[2]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Radiology, İstanbul, Turkey
[3]Medicana Ataköy Hospital, Clinic of Radiology, İstanbul, Turkey
[4]Memorial Göztepe Hospital, Clinic of Radiology, İstanbul, Turkey

Abstract

Objective: The aim of the current study was to investigate factors associated with recurrent bleeding following bronchial artery embolization (BAE), with particular emphasis on the role of preprocedural computed tomography angiography (CTA).

Method: This retrospective study included 110 patients who underwent BAE for hemoptysis. Patients with and without recurrent bleeding were compared for clinical characteristics, underlying etiology, angiographic features including bronchial and non-bronchial systemic arterial involvement, the number and type of embolized vessels, as well as procedural parameters such as fluoroscopy time, procedure duration, and radiation dose. Factors associated with recurrent bleeding were evaluated using univariate analyses followed by multivariable logistic regression.

Results: Recurrent bleeding occurred in 25 patients (22.7%). The incidence of recurrence was significantly higher in patients without preprocedural CTA compared to those who underwent CTA (36.5% vs. 10.3%, p=0.001). Multivariate analysis demonstrated that the lack of preprocedural CTA was independently associated with recurrent bleeding (odds ratio: 4.985; 95% confidence interval: 1.784-13.926; p=0.002). Mortality was also higher among patients with recurrent however, the difference was not statistically significant (36.0% vs. 21.2% p=0.130).

Öz

Amaç: Bu çalışmanın amacı, bronşiyal arter embolizasyonu (BAE) sonrası rekürren kanama ile ilişkili faktörleri araştırmak ve özellikle işlem öncesi bilgisayarlı tomografi anjiyografinin (BTA) rolüne odaklanmaktır.

Yöntem: Bu retrospektif çalışmaya hemoptizi nedeniyle bronşiyal arter embolizasyonu uygulanan 110 hasta dahil edildi. Rekürren kanama gelişen ve gelişmeyen hastalar; klinik özellikler, altta yatan etiyolojiler, bronşiyal ve bronşiyal olmayan sistemik arter tutulumu, embolize edilen damar sayısı ve tipini içeren anjiyografik bulgular ile floroskopi süresi, işlem süresi ve radyasyon dozu gibi işlem parametreleri açısından karşılaştırıldı. Rekürren kanama ile ilişkili faktörler önce tek değişkenli analizler ile, ardından çok değişkenli lojistik regresyon analizi ile değerlendirildi.

Bulgular: Rekürren kanama 25 hastada (%22,7) gözlendi. İşlem öncesi BTA mevcut olmayan hastalarda rekürrens insidansı, BTA yapılan hastalara kıyasla anlamlı derecede daha yüksekti (%36,5'e karşı %10,3; p=0,001). Çok değişkenli analizde, işlem öncesi BTA'nın olmaması rekürren kanama ile bağımsız olarak ilişkili bulundu (odds oranı: 4,985; %95 güven aralığı: 1,784-13,926; p=0,002). Mortalite sayısal olarak daha yüksekti; ancak fark istatistiksel olarak anlamlı değildi (%36,0'a karşı %21,2; p=0,130).

Address for Correspondence: Eda Cingöz, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Radiology, İstanbul, Turkey
E-mail: edacanipek@gmail.com **ORCID:** orcid.org/0000-0003-0814-4597
Received: 03.05.2026 **Accepted:** 03.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Cingöz E, Dablan A, Arslan MF, Türksayar O, Cingöz M, Erdim Ç, et al. Factors associated with the recurrence of hemoptysis following bronchial artery embolization. Bagcilar Med Bull. 2026;11(3):350-358

Abstract

Conclusion: The availability of preprocedural CTA prior to embolization is associated with lower rates of recurrent bleeding. It may facilitate a more complete identification of culprit vessels and reduce the likelihood of missed sources and subsequent recurrence.

Keywords: Bronchial artery embolization, computed tomography angiography, hemoptysis, recurrence

Öz

Sonuç: İşlem öncesi BTA'nın mevcut olması, bronşiyal arter embolizasyonu sonrası daha düşük rekürren kanama oranları ile ilişkilidir. BTA, sorumlu damarların daha kapsamlı tanımlanmasını sağlayarak gözden kaçan kanama kaynaklarına bağlı rekürrens olasılığının azaltılmasına katkıda bulunabilir.

Anahtar kelimeler: Bilgisayarlı tomografi anjiyografi, bronşiyal arter embolizasyonu, hemoptizi, rekürrens

Introduction

Hemoptysis is a potentially serious condition that may range from spontaneously resolving bleeding to a fatal compromise of the airways and hemodynamic instability. The underlying etiology varies, depending on geographic and clinical factors; however, lung cancer, active tuberculosis and its sequelae, bronchiectasis, and aspergillosis account for the majority (around 80%) of the cases (1,2).

Severe hemoptysis is considered to be a respiratory emergency and, if left untreated, may be associated with mortality rates as high as 50-100%, most commonly due to asphyxiation rather than exsanguination. Nonetheless, even minor initial episodes, such as blood-streaked sputum or low-volume "index bleeding", may precede life-threatening hemoptysis and therefore warrant early recognition and intervention (3-5). Bronchoscopy plays an important role in maintaining airway patency and localizing the source of bleeding, while surgery remains a definitive option in selected patients; however, both approaches may be limited in an acute setting or in patients at a high risk of mortality (6,7). Bronchial artery embolization (BAE), first introduced into clinical practice in 1973, has evolved into a widely used, minimally invasive, and effective treatment option for controlling hemoptysis. Moreover, advances in microcatheter technology and embolic materials have improved procedural safety and outcomes (8). Despite high technical success and effective initial bleeding control, recurrent bleeding after BAE remains a significant clinical challenge, with reported recurrence rates of up to 57% and an association with increased mortality (9-13). Several factors have been implicated in hemoptysis recurrence after BAE, including incomplete embolization, involvement of non-bronchial systemic arteries (NBSA), larger vessel caliber, as well as recanalization and neovascularization. Given the relatively high recurrence rates, a systematic evaluation of multiple potentially contributing factors is warranted to better understand recurrence patterns and guide clinical management.

The aim of the current study was to investigate factors associated with the recurrence of hemoptysis following BAE, with a particular focus on clinical, angiographic, and procedural variables.

Materials and Methods

Approval for this retrospective study was obtained from the University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital Clinical Research Ethics Committee (approval number: 2024-131, date: 24.07.2024), and owing to the retrospective design, the requirement for informed consent was waived.

A total of 111 consecutive patients presenting with hemoptysis at the departments of pulmonology or thoracic surgery and subsequently referred to interventional radiology for BAE were retrospectively evaluated between January 2021 and January 2022. One patient was excluded due to vascular anatomical features that precluded completion of the procedure, leaving a final study cohort of 110 patients who were followed for 12-24 months. Imaging data were retrieved from the institutional picture archiving and communication system and clinical data were retrieved from the electronic medical record system.

The clinical variables evaluated were age, sex, comorbidities, and the presence, frequency, and timing of recurrent hemoptysis following the initial procedure, along with preprocedural computed tomography angiography (CTA), if available. Angiographic and procedural variables included the number and type of embolized bronchial arteries, the presence, location and number of NBSA embolizations, the total number of embolized vessels, fluoroscopy time, procedure duration, and total air kerma. Additionally, angiographic features on digital subtraction angiography, including hypervascularity and tortuosity, and the embolic agents used were also recorded. BAE

procedures were carried out by a team comprising seven interventional radiologists, each with a minimum of five years of experience. The clinical, imaging, and procedural variables collected were compared between patients with and without recurrent hemoptysis.

All procedures were carried out under local anesthesia. Access to the vascular system was obtained through the common femoral artery under ultrasound guidance. Angiography targeting the bronchial arteries and NBSAs thought to be responsible for bleeding was carried out using a 5-F catheter (Cobra or Simmons-1), and additional arteriograms of the bronchial, intercostal, and subclavian territories were obtained to identify abnormal contrast opacification. A 5-F vertebral or Berenstein catheter was used for evaluation of the subclavian artery when indicated. Preprocedural CTA was evaluated, where available, for NBSA feeders and to delineate all potential culprit vascular sources contributing to the hemoptysis. Angiographic runs were carefully reviewed to identify and avoid branches supplying vital tissues, particularly the spinal cord, and heightened attention was applied to vessels originating from the aortic arch, the internal mammary artery, and branches of the subclavian artery to limit unintended embolization. Common bronchial arterial trunks and the left bronchial artery were accessed selectively with standard devices and treated in the absence of retrograde flow, whereas vessels arising from the intercostobronchial trunk, the intercostal circulation, branches of the subclavian system, or the phrenic circulation were managed using a highly selective approach via a coaxial microcatheter system (2.4-2.7 F; Terumo, Tokyo, Japan) in order to decrease off-target embolization. Embolization was performed using polyvinyl alcohol particles (Contour; Boston Scientific, Natick, MA, USA), measuring 500-1180 µm, until adequate flow stasis was achieved. The procedure was concluded once full embolization of all culprit bronchial arteries and NBSAs had been secured (Figures 1, 2). Adjunctive coil and/or n-butyl cyanoacrylate (NBCA)–lipiodol mixture embolization was applied in selected cases based on the preference of the operator.

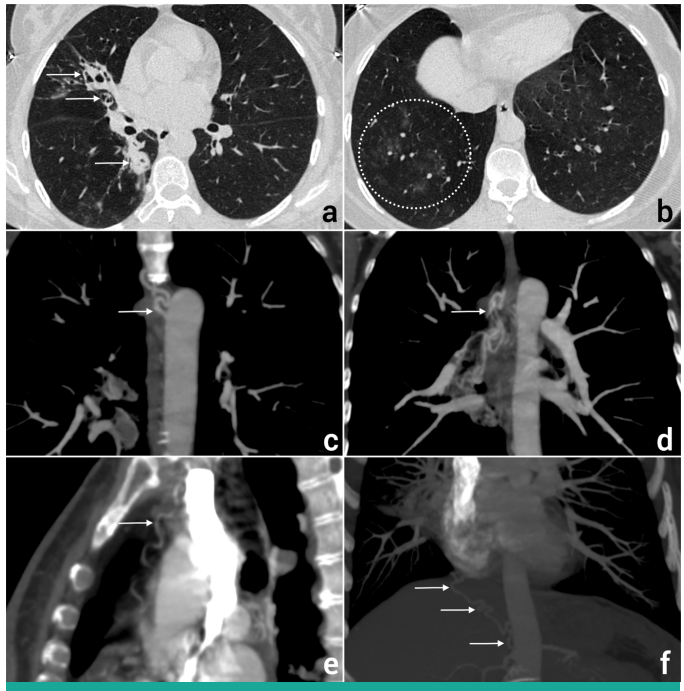

Figure 1. Images of a 61-year-old female patient presenting with hemoptysis, who underwent preprocedural computed tomography (CT) angiography and was subsequently referred to interventional radiology for bronchial artery embolization. a) Axial thorax CT image shows right lung bronchiectasis (white arrows). b) A more inferior axial section demonstrates patchy ground-glass opacities consistent with alveolar hemorrhage (broken white circle) in the patient presenting with hemoptysis. c, d) Coronal maximum intensity projection (MIP) images show hypertrophied and tortuous right intercostobronchial and right bronchial arteries, respectively. e, f) Sagittal and coronal MIP images demonstrate hypertrophied and tortuous right internal mammary and right phrenic artery branches, respectively

Statistical Analysis

Statistical analyses were carried out via IBM SPSS software version 31. The distribution of continuous variables was evaluated with the Shapiro-Wilk test, along with graphical methods. Variables not following a normal distribution were expressed as median (Q1-Q3), while categorical data were presented as counts and percentages. When two groups were compared, the Mann-Whitney U analysis was employed for continuous data that were not normally distributed, while the Pearson chi-square method was utilized for categorical data. Factors associated with recurrent bleeding were first evaluated using univariate analyses. Fisher's exact test was used for 2×2 contingency tables when the assumptions of Pearson's chi-square test were not met, and all statistical tests were two-sided. Parameters showing statistical significance in univariate analyses, together with those deemed clinically meaningful, were then entered into a multivariate model. Etiological variables were not incorporated into the multivariate model because of the small sample size and concerns regarding overfitting. A logistic regression model was employed to determine independent factors associated with recurrent bleeding, with results reported as odds ratios (ORs) and corresponding 95% confidence intervals (CIs). Model calibration was evaluated with the Hosmer-

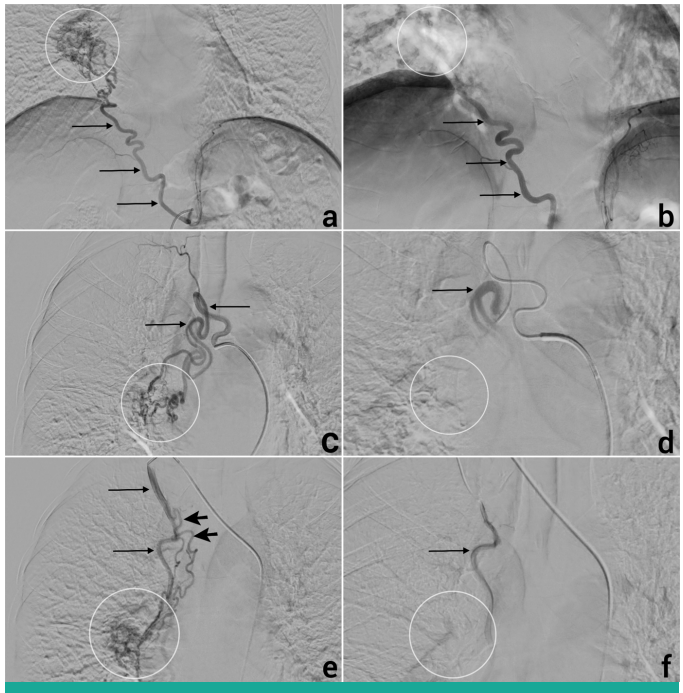

Figure 2. Digital subtraction angiography (DSA) of the patient from Figure 1 who subsequently underwent bronchial artery embolization at the interventional radiology unit. a) Anteroposterior (AP) DSA image demonstrates a focal hypervascularity and vascular blush showing the presumed bleeding site (white circle), that was supplied by the right phrenic artery branches (black arrows). b) Subsequent AP DSA following selective microcatheter embolization shows that the area in a could no longer be visualized (white circle), with flow stasis in the right phrenic artery (black arrows). c) AP DSA after intercostobronchial artery catheterization demonstrates an area of abnormal hypervascularity (white circle) consistent with the suspected bleeding source arising from a right bronchial branch of the intercostobronchial trunk (black arrows). d) AP DSA after selective embolization shows resolution of the area in c (white circle), with proximal stasis in the right bronchial artery (black arrow). e) AP DSA of the right internal mammary artery demonstrates a culprit vascular territory (white circle) supplied by its branches (short arrows) from the main trunk (long arrows). f) Subsequent AP DSA following selective embolization shows disappearance of the area in e (white circle) while preserving patency of the right internal mammary artery (black arrow)

Lemeshow test, while model performance was quantified using the Nagelkerke R^2 coefficient. Statistical significance was defined as a p-value below 0.05.

Results

The study cohort comprised 110 patients. Recurrent bleeding occurred in 25 patients (22.7%) whereas 85 patients (77.3%) had no recurrence. The median age was 57.5 years (interquartile range 44.5-64.2), with a predominance of male patients (n=83, 75.5%) compared to females (27, 24.5%). The most common underlying conditions were pneumonia (20.9%), tuberculosis (16.4%), bronchiectasis (16.4%), and aspergilloma (15.5%), while less frequent comorbidities included chronic obstructive pulmonary disease (5.5%), cardiac pathologies (4.5%), and other diseases. Mortality analysis showed that 27 patients (24.5%) died, whereas 83 patients (75.5%) survived. Preprocedural CTA was available for 58 patients (52.7%), while 52 patients (47.3%) underwent BAE without any preprocedural CTA evaluation.

Age (p=0.889) and sex distribution (p=0.648) did not differ significantly between patients with and without recurrent bleeding. The underlying etiologies differed between the groups, with pneumonia (21.2%) and bronchiectasis (21.2%) being the most frequent in patients without recurrence, while tuberculosis (36.0%) and aspergilloma (24.0%) were predominant among the patients who developed recurrent bleeding. Patients with recurrent bleeding showed a higher mortality rate (36.0%) compared with those without recurrence (21.2%). A significantly higher rate of recurrent bleeding was observed among patients who had not undergone preprocedural CTA (36.5%) compared with those who had (10.3%) (p=0.001; Table 1). A multivariate logistic regression model including age, sex, and preprocedural CTA was statistically significant overall (Omnibus test, p=0.011). The lack of preprocedural CTA was independently associated with recurrent bleeding (OR: 4.985; 95% CI: 1.784-13.926; p=0.002). The lack of preprocedural CTA was independently associated with recurrent bleeding (OR: 4.985; 95% CI: 1.784-13.926; p=0.002), demonstrating a remarkable fivefold increase in the odds of recurrence. Neither age nor sex demonstrated a significant association with recurrent bleeding (p>0.05) (Table 2).

The median time to recurrent bleeding was 2 months (1-5.5), and the median number of recurrent bleeding episodes was 2 (1-3). No statistically significant differences were identified in the number of embolized bronchial arteries between the patients with and without recurrent bleeding (p=0.678). Evaluation of the types of embolized bronchial arteries, including the right bronchial artery, left bronchial artery, common trunk of bronchial arteries, and right intercostobronchial trunk, indicated no significant differences between the two groups (p=0.764, p=0.884, p=0.693, and p=0.677, respectively). Similarly, the use of non-bronchial embolization was comparable between patients with and without recurrence (p=0.891). In addition, neither the number of non-bronchial arteries embolized

Table 1. Comparison of demographic and clinical characteristics of patients with and without recurrent bleeding

Variables	All patients (n=110)	Recurrent bleeding		p
		Absent (n=85)	Present (n=25)	
Age/median (1st-3rd quartiles)	57.5 (44.5-64.2)	58.0 (43.0-63.0)	56.0 (46.5-66.0)	0.889[a]
Sex/n (%)				
Male	83 (75.5)	65 (76.5)	18 (72.0)	0.648[b]
Female	27 (24.5)	20 (23.5)	7 (28.0)	
Comorbidity*/n (%)				
Malignancy	1 (0.9)	1 (1.2)	0	
Tuberculosis	18 (16.4)	9 (10.6)	9 (36.0)	
Pneumonia	23 (20.9)	18 (21.2)	5 (20.0)	
Bronchiectasis	18 (16.4)	18 (21.2)	0	
Aspergilloma	17 (15.5)	11 (12.9)	6 (24.0)	
APVR	1 (0.9)	1 (1.2)	0	
Cardiac disease	5 (4.5)	5 (5.9)	0	
Chronic granulomatous disease	2 (1.8)	2 (2.4)	0	
Pulmonary embolism	1 (0.9)	1 (1.2)	0	
Asthma	3 (2.7)	2 (2.4)	1 (4.0)	
Pulmonary artery aneurysm	1 (0.9)	1 (1.2)	0	
Pneumothorax	2 (1.8)	2 (2.4)	0	
COPD	6 (5.5)	5 (5.9)	1 (4.0)	
Churg-Strauss	3 (2.7)	2 (2.4)	0	
Hypertension	1 (0.9)	1 (1.2)	0	
Behçet disease	1 (0.9)	1 (1.2)	0	
HIV positivity	1 (0.9)	1 (1.2)	0	
Malaria	1 (0.9)	1 (1.2)	0	
Tuberous sclerosis	1 (0.9)	1 (1.2)	0	
Mortality/n (%)				
Deceased	27 (24.5)	18 (21.2)	9 (36.0)	0.130[b]
Alive	83 (75.5)	67 (78.8)	16 (64.0)	
Preprocedural CTA/n (%)				
Absent	52 (47.3)	33 (63.5)	19 (36.5)	0.001[b]
Present	58 (52.7)	52 (89.7)	6 (10.3)	

[a]: Mann-Whitney U test, [b]: Pearson chi-square test, *: Some patients had more than one comorbidity.
APVR: Anomalous pulmonary venous return, HIV: Human immunodeficiency virus, COPD: Chronic obstructive pulmonary disease, CTA: Computed tomography angiography, percentages for preprocedural CTA are expressed as row percentages

Table 2. Multivariable logistic regression analysis of factors associated with recurrent bleeding

Variable	B	OR [Exp(B)]	95% CI	p
Age	0.000	1.000	0.967-1.035	0.992
Sex (male)	0.006	1.006	0.332-3.052	0.991
Preprocedural CTA (absent)	1.606	4.985	1.784-13.926	0.002

Omnibus test: p=0.011; Nagelkerke R^2: 0.145; Hosmer-Lemeshow: p=0.672
Dependent variable: Recurrent bleeding (0= no, 1= yes)
OR: Odds ratio, CI: Confidence interval, CTA: Computed tomography angiography

nor the total number of embolized vessels showed a significant difference between the groups (p=0.474 and p=0.751, respectively) (Table 3).

Fluoroscopy time, procedure duration, and total air kerma did not differ significantly between the patients with and without recurrent bleeding (p=0.225, p=0.283, and p=0.160, respectively). Similarly, no significant differences were observed in the presence of hypervascularity or tortuosity between the two groups of patients (p=1.000 and p=0.468, respectively). Among the 66 patients with available embolic-agent data, particulate embolic agents were used in all cases. The use of adjunctive glue and coil embolization was comparable between the two groups (p=0.318 and p=1.000, respectively) (Table 4).

Evaluation of the procedural parameters according to the availability of preprocedural CTA indicated no significant differences in fluoroscopy time, procedure duration, or total air kerma in the cohort overall (p=0.695, p=0.801, and p=0.923, respectively). Similarly, these parameters remained comparable irrespective of CTA status among the patients without any recurrent bleeding (p=0.839, p=0.815, and p=0.849, respectively) (Table 5).

Discussion

In this study, lack of preprocedural CTA emerged as an independent factor associated with increased odds of recurrent bleeding, and recurrent bleeding tended to be associated with higher mortality.

BAE is a widely applied treatment for hemoptysis with relatively few limitations. Absolute contraindications are mainly related to vessels supplying blood to critical structures such as the heart, brain, or spinal cord; however, in most cases, the risk associated with untreated massive hemoptysis outweighs the procedural risks (3). Despite high initial success rates, recurrent bleeding remains a significant clinical challenge (9,13,14). Early recurrence is often related to incomplete embolization or failure to identify all bleeding sources, whereas late recurrence is typically associated with disease progression, recanalization, or the development of collateral circulation (15). Overlooked culprit vessels have been identified as a leading cause of recurrent hemoptysis on repeat angiography (14).

Efforts to reduce recurrence following BAE have mainly focused on improving the detection and embolization of all

Table 3. Comparison of embolization and angiographic characteristics according to recurrent bleeding

Variables	Recurrent bleeding		p
	Absent (n=85)	Present (n=25)	
Time to recurrent bleeding (months)/median (Q1-Q3)	-	2 (1-5.5)	
Number of recurrent bleeding episodes/median (Q1-Q3)	-	2 (1-3)	
Number of embolized bronchial arteries/median (Q1-Q3)	2 (1-2)	2 (1-2)	0.678[b]
Distribution of embolized arteries*/n (%)			
Right bronchial artery	18 (21.2)	6 (24.0)	0.764[a]
Left bronchial artery	36 (42.4)	11 (44.0)	0.884[a]
Common trunk of bronchial arteries	48 (56.5)	13 (52.0)	0.693 [a]
Right intercostobronchial artery	47 (55.3)	15 (60.0)	0.677[a]
Non-bronchial systemic artery embolization/n (%)	25 (29.4)	7 (28.0)	0.891[a]
Distribution of non-bronchial systemic arteries embolized (n=32)*/n (%)			
Left internal mammary artery	13 (52.0)	4 (57.1)	1.000[c]
Left subclavian artery	1 (4.0)	-	-
Left phrenic artery	2 (8.0)	-	-
Right internal mammary artery	6 (24.0)	2 (28.6)	1.000[c]
Right subclavian artery	2 (8.0)	-	
Right phrenic artery	2 (8.0)	-	
Aberrant/accessory bronchial artery	3 (12.0)	-	
Intercostal artery	7 (28.0)	2 (28.6)	1.000[c]
Subclavian branches	3 (12.0)	1 (14.3)	1.000[c]
Number of non-bronchial systemic arteries embolized (n=32)/median (Q1-Q3)	1 (1-2)	1 (1-2)	0.474[b]
Total number of embolized vessels/median (Q1-Q3)	2 (1.5-3)	2 (2-3)	0.751[b]

[a]: Chi-square test, [b]: Pearson chi-square test; [c]: Fisher's exact test (two-sided), *: Some patients underwent embolization of more than one vessel

Table 4. Comparison of procedural parameters, angiographic findings, and embolic agents used in patients with and without recurrent bleeding

Variables	Recurrent bleeding		p
	Absent (n=85)	Present (n=25)	
Fluoroscopy time (min)/median (Q1-Q3)	27.3 (19.1-35.0)	24.0 (19.0-30.0)	0.225[a]
Procedure duration (min)/median (Q1-Q3)	41.0 (29.0-54.5)	37.0 (26.5-45.0)	0.283[a]
Total air kerma (mGy)/median (Q1-Q3)	515.0 (302.0-961.0)	483.0 (262.0-561.0)	0.160[a]
Hypervascularity/n (%)			
Absent	11 (12.9)	3 (12.0)	1.000[c]
Present	74 (87.1)	22 (88.0)	
Tortuosity/n (%)			
Absent	41 (48.2)	10 (40.0)	0.468[b]
Present	44 (51.8)	15 (60.0)	
Type of embolic agent used *(n=66)/n (%)			
Particle	51 (100.0)	15 (100.0)	-
Glue	3 (5.9)	2 (13.3)	0.318[c]
Coil	10 (19.6)	3 (20.0)	1.000[c]

[a]: Mann-Whitney U test, [b]: Pearson chi-square test; [c]: Fisher's exact test (two-sided) *: Analyses of embolic agents were performed only in patients with available data (n=66). More than one embolic agent was used in some patients

Table 5. Comparison of procedural parameters according to preprocedural CTA availability stratified by recurrence status

Group	CTA	Fluoroscopy time (min)	Procedure time (min)	Total air kerma (mGy)
		Median (1st-3rd quartile)	Median (1st-3rd quartile)	Median (1st-3rd quartile)
Entire cohort (n=110)	Absent	27.2 (21.2-31.6)	39.0 (29.0-49.2)	486.0 (295.2-878.0)
	Present	26.9 (17.0-33.5)	40.0 (23.5-52.7)	520.0 (263.0-714.2)
	p*	0.695	0.801	0.923
Patients without recurrent bleeding (n=85)	Absent	27.9 (21.4-33.9)	41.0 (29.5-54.0)	512.0 (282.0-1021.0)
	Present	27.0 (17.5-35.0)	40.5 (21.2-57.2)	520.0 (298.2-758.7)
	p*	0.839	0.815	0.849
Patients with recurrent bleeding (n=25)	Absent	25.0 (20.0-30.0)	38.0 (27.0-45.0)	483.0 (295.5-611.0)
	Present	19.6 (14.0-24.2)	31.5 (21.0-43.0)	408.5 (128.2-710.5)
	p*	0.092	0.366	0.708

*: Mann-Whitney U test, CTA: Computed tomography angiography

potential bleeding sources, specifically ectopic bronchial arteries and NBSAs, which are frequently implicated in recurrent hemoptysis. Conventional angiography may fail to identify all relevant vascular contributors, particularly small-caliber or tortuous vessels and those arising from atypical or extra-thoracic origins, which may limit comprehensive assessment of the source of bleeding. These limitations have led to increased interest in preprocedural CTA as an adjunctive imaging modality to improve vascular mapping prior to embolization.

Preprocedural CTA has been reported to achieve high diagnostic accuracy, approaching 97.5%, in identifying culprit bronchial arteries and NBSA, thereby facilitating more comprehensive and targeted embolization (16). In addition, it provides superior delineation of the vascular anatomy, particularly in identifying NBSA and anatomic variants that may be overlooked on conventional angiography (17). Furthermore, CTA plays an important role in procedural planning by enabling precise identification of target vessels, which may help streamline the catheterization (18). Consistent with these findings, CTA has also been reported to facilitate the detection of ectopic bronchial arteries and NBSA, particularly those originating from the subclavian and internal mammary arteries, and has been associated with improved hemoptysis-free early survival, supporting its routine use prior to BAE (19).

A prospective study by Le et al. (16), which included 57 patients, reported high diagnostic accuracy (97.5%)

in the identification of culprit arteries with the use of multidetector CTA and was associated with relatively low early recurrence rates (16.7%). In our cohort of 110 patients, recurrent bleeding occurred in 22.7% of cases, with tuberculosis and aspergilloma being the most frequent underlying conditions among the patients with recurrence (36.0% and 24.0%, respectively). Recurrent bleeding in our cohort demonstrated a trend, albeit non-significant, toward increased mortality, consistent with prior reports (13,15). Failure to reach statistical significance may be related to the small sample size. Accurate identification and complete treatment of all potential bleeding sources are therefore essential, which in turn depends on a thorough preprocedural evaluation of the vascular anatomy. In line with prior studies suggesting the clinical relevance of comprehensive vascular mapping and early outcome benefits associated with preprocedural imaging, the present study demonstrates that the lack of preprocedural CTA was independently associated with an increased risk of recurrent bleeding following BAE (16,19).

An important point to consider is that preprocedural CTA was not assigned in a randomized manner in the current study, and the decision to perform CTA may have been influenced by clinical presentation, availability, or clinician preference. Therefore, the potential for selection bias and unmeasured confounding cannot be excluded. It is possible that patients with a more complex disease or challenging vascular anatomy may have been more likely to undergo CTA. Alternatively, patients receiving more comprehensive evaluation and management may have been preferentially selected for CTA. These factors may have contributed, at least in part, to the observed association between the absence of CTA and increased recurrence.

No significant differences were observed in total air kerma, fluoroscopy time, or overall procedural duration between the patients with and without recurrent bleeding. This finding may be partly explained by the additional time required for the embolization of additional vascular territories identified on preprocedural CTA, which could have mitigated potential differences in procedural efficiency between the groups. Moreover, variations in operator experience and technique likely introduced a degree of heterogeneity, further contributing to the overlap in procedural metrics. Taken together, these factors may have negated any measurable differences in exposure to radiation and procedure duration between the patient groups.

Study Limitations

A number of limitations should be taken into account when interpreting the findings. First, the follow-up period was 12-24 months; this duration largely covers the interval during which hemoptysis recurrence after BAE is most frequently observed in the literature. However, within this cohort, the median interval to recurrence was 2 months. Moreover, the relatively limited sample size (n=110) may have constrained the statistical power to identify certain associations. Second, the single-center design may restrict the generalizability of our findings, as patient characteristics, procedural approaches, and operator experience may vary across institutions. In addition, hemoptysis recurrence is a multifactorial process that is influenced by the underlying etiology, vascular anatomy, and procedural factors; therefore, the results warrant cautious interpretation and should not be directly extrapolated to all patient populations. Further prospective, preferably randomized, studies are warranted to better define the factors associated with recurrent hemoptysis following BAE.

Conclusion

Preprocedural CTA may guide the optimal plan for BAE by enabling comprehensive delineation of bronchial and NBSA supply, including accessory and ectopic vessels. Such a detailed anatomic assessment can facilitate a more targeted embolization strategy, reducing the likelihood of missed culprit vessels and subsequent recurrent bleeding. Accordingly, the incorporation of CTA into the preprocedural evaluation may contribute towards improved clinical outcomes in patients undergoing embolization for hemoptysis.

Ethics

Ethics Committee Approval: Approval for this retrospective study was obtained from the University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital Clinical Research Ethics Committee (approval number: 2024-131, date: 24.07.2024). The study was conducted in accordance with the Declaration of Helsinki.

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

Footnotes

Authorship Contributions

Concept: E.C., A.D., Ö.K., Design: A.D., M.C., Ç.E., İ.N.M., Ö.K., Data Collection or Processing: M.F.A., O.T., İ.N.M.,

T.G., Analysis or Interpretation: A.D., M.F.A., M.C., T.G., Ö.K., Literature Search: E.C., M.F.A., O.T., Ç.E., İ.N.M., Writing: E.C., O.T., M.C., Ç.E., T.G.

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

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

References

1. Fartoukh M, Khalil A, Louis L, Carette MF, Bazelly B, Cadranel J, et al. An integrated approach to diagnosis and management of severe haemoptysis in patients admitted to the intensive care unit: a case series from a referral centre. Respir Res. 2007;8(1):11.
2. Abdulmalak C, Cottenet J, Beltramo G, Georges M, Camus P, Bonniaud P, et al. Haemoptysis in adults: a 5-year study using the French nationwide hospital administrative database. Eur Respir J. 2015;46(2):503-511.
3. Kettenbach J, Ittrich H, Gaubert JY, Gebauer B, Vos JA. CIRSE standards of practice on bronchial artery embolisation. Cardiovasc Intervent Radiol. 2022;45(6):721-732.
4. Parrot A, Tavolaro S, Voiriot G, Canellas A, Assouad J, Cadranel J, et al. Management of severe hemoptysis. Expert Rev Respir Med. 2018;12(10):817-829.
5. Sidhu M, Wieseler K, Burdick TR, Shaw DW. Bronchial artery embolization for hemoptysis. Semin Intervent Radiol. 2008;25(3):310-318.
6. Davidson K, Shojaee S. Managing massive hemoptysis. Chest. 2020;157(1):77-88.
7. Radchenko C, Alraiyes AH, Shojaee S. A systematic approach to the management of massive hemoptysis. J Thorac Dis. 2017;9(Suppl 10):S1069-S1086.
8. Rémy J, Voisin C, Dupuis C, Beguery P, Tonnel AB, Denies JL, et al. [Treatment of hemoptysis by embolization of the systemic circulation]. Ann Radiol (Paris). 1974;17(1):5-16.
9. Li H, Ding X, Zhai S, Gao K. A retrospective study on the management of massive hemoptysis by bronchial artery embolization: risk factors associated with recurrence of hemoptysis. BMC Pulm Med. 2023;23(1):87.
10. van den Heuvel MM, Els Z, Koegelenberg CF, Naidu KM, Bolliger CT, Diacon AH. Risk factors for recurrence of haemoptysis following bronchial artery embolisation for life-threatening haemoptysis. Int J Tuberc Lung Dis. 2007;11(8):909-914.
11. Dorji K, Hongsakul K, Jutidamrongphan W, Oofuvong M, Geater S. Bronchial artery embolization in life-threatening hemoptysis: outcome and predictive factors. J Belg Soc Radiol. 2021;105(1):5.
12. Ishikawa H, Ohbe H, Omachi N, Morita K, Yasunaga H. Spinal cord infarction after bronchial artery embolization for hemoptysis: a nationwide observational study in Japan. Radiology. 2021;298(3):673-679.
13. Panda A, Bhalla AS, Goyal A. Bronchial artery embolization in hemoptysis: a systematic review. Diagn Interv Radiol. 2017;23(4):307-317.
14. Xu S, Guan LJ, Shi BQ, Tan YS, Zhang XJ. Recurrent hemoptysis after bronchial artery embolization: prediction using a nomogram and artificial neural network model. AJR Am J Roentgenol. 2020;215(6):1490-1498.
15. Yoon W, Kim JK, Kim YH, Chung TW, Kang HK. Bronchial and nonbronchial systemic artery embolization for life-threatening hemoptysis: a comprehensive review. Radiographics. 2002;22(6):1395-1409.
16. Le HY, Le VN, Pham NH, Phung AT, Nguyen TT, Do Q. Value of multidetector computed tomography angiography before bronchial artery embolization in hemoptysis management and early recurrence prediction: a prospective study. BMC Pulm Med. 2020;20(1):231.
17. Ravetta P, Vouche M. Added value of computed tomography angiography prior to bronchial artery embolization for hemoptysis: a retrospective two-center study. J Belg Soc Radiol. 2023;107(1):00.
18. Singhal M, Lal A, Prabhakar N, Yadav MK, Vijayvergiya R, Behra D, et al. Non-bronchial causes of haemoptysis: imaging and interventions. Pol J Radiol. 2020;85:e328-e339.
19. Li PJ, Yu H, Wang Y, Jiang FM, Wang W, Li XO, et al. Multidetector computed tomography angiography prior to bronchial artery embolization helps detect culprit ectopic bronchial arteries and non-bronchial systemic arteries originating from subclavian and internal mammary arteries and improve hemoptysis-free early survival rate in patients with hemoptysis. Eur Radiol. 2019;29(4):1950-1958.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):359-365
DOI: 10.4274/BMB.galenos.2026.46036

Prognostic Impact of Combined Anemia and Hypoalbuminemia Phenotypes in Patients Hospitalized with Acute Heart Failure

Akut Kalp Yetersizliği Hastalarında Kombine Anemi ve Hipoalbüminemi Fenotiplerinin Prognostik Etkisi

Can Baba Arın[1], Mehmet Zafer Aydın[2], Mehmet Kamil Teber[2], Nizamettin Selçuk Yelgeç[1]

[1]University of Health Sciences Turkey, Dr. Siyami Ersek Thoracic and Cardiovascular Surgery Training and Research Hospital, Department of Cardiology, İstanbul, Turkey
[2]University of Health Sciences Turkey, Kayseri City Training and Research Hospital, Department of Cardiology, Kayseri, Turkey

Abstract

Objective: Hematologic and nutritional abnormalities are common in patients hospitalized with acute heart failure and may be associated with adverse outcomes. This study aimed to evaluate the clinical characteristics and in-hospital mortality of patients classified according to combined anemia and hypoalbuminemia phenotypes.

Method: This retrospective study included 220 consecutive patients hospitalized with acute heart failure. Anemia was defined according to World Health Organization sex-specific hemoglobin thresholds, and hypoalbuminemia was defined as a serum albumin level <3.5 g/dL. Patients were classified into four phenotypes: Neither anemia nor hypoalbuminemia, anemia alone, hypoalbuminemia alone, and both anemia and hypoalbuminemia. Associations between phenotype groups and all-cause in-hospital mortality were evaluated using categorical comparisons and logistic regression. Given the limited number of mortality events, a parsimonious age-adjusted model and bootstrap sensitivity analysis were performed.

Results: Among 220 patients, 21 (9.5%) died during hospitalization. The four phenotype groups comprised 39 patients with neither abnormality, 34 with anemia alone, 45 with hypoalbuminemia alone, and 102 with both abnormalities. In-hospital mortality rates were 7.7%, 17.6%, 8.9%, and 7.8%, respectively, with no significant overall difference among groups (Fisher's exact p=0.402). In the age-adjusted logistic regression model, phenotype group was not significantly

Öz

Amaç: Akut kalp yetersizliği nedeniyle hastaneye yatırılan hastalarda hematolojik ve nütrisyonel bozukluklar sık görülmekte olup olumsuz klinik sonuçlarla ilişkili olabilir. Bu çalışmada, kombine anemi-hipoalbüminemi fenotiplerine göre sınıflandırılan hastaların klinik özelliklerinin ve hastane içi mortalitelerinin değerlendirilmesi amaçlandı.

Yöntem: Bu retrospektif çalışmaya akut kalp yetersizliği nedeniyle hastaneye yatırılan 220 ardışık hasta dahil edildi. Anemi, Dünya Sağlık Örgütü'nün cinsiyete özgü hemoglobin eşik değerlerine göre; hipoalbüminemi ise serum albümin düzeyinin <3,5 g/dL olması olarak tanımlandı. Hastalar anemi ve hipoalbüminemi bulunmayanlar, yalnızca anemisi olanlar, yalnızca hipoalbüminemisi olanlar ve hem anemi hem de hipoalbüminemisi olanlar olmak üzere dört fenotipe ayrıldı. Fenotip grupları ile tüm nedenlere bağlı hastane içi mortalite arasındaki ilişkiler kategorik karşılaştırmalar ve lojistik regresyon analizleriyle değerlendirildi. Mortalitenin sınırlı sayıda olması nedeniyle parsimoniyöz, yaşa göre düzeltilmiş bir model ve bootstrap duyarlılık analizi uygulandı.

Bulgular: Toplam 220 hastanın 21'i (%9,5) hastanede yatış sırasında yaşamını yitirdi. Anemi ve hipoalbüminemisi olmayan grupta 39, yalnızca anemisi olan grupta 34, yalnızca hipoalbüminemisi olan grupta 45 ve her iki bozukluğun birlikte bulunduğu grupta 102 hasta vardı. Hastane içi mortalite oranları sırasıyla %7,7, %17,6, %8,9 ve

Address for Correspondence: Can Baba Arın, University of Health Sciences Turkey, Dr. Siyami Ersek Thoracic and Cardiovascular Surgery Training and Research Hospital, Department of Cardiology, İstanbul, Turkey
E-mail: canbabaarin@gmail.com **ORCID:** orcid.org/0000-0003-1647-3120
Received: 16.05.2026 **Accepted:** 23.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Arın CB, Aydın MZ, Teber MK, Yelgeç NS. Prognostic impact of combined anemia and hypoalbuminemia phenotypes in patients hospitalized with acute heart failure. Bagcilar Med Bull. 2026;11(3):359-365

Abstract

associated with in-hospital mortality (overall p=0.391), and none of the individual phenotype comparisons reached statistical significance. Bootstrap sensitivity analysis yielded consistent findings. Receiver operating characteristic analysis showed limited discrimination for mortality for both albumin [area under the curve (AUC) 0.594, 95% confidence interval (CI) 0.477-0.711; p=0.157] and hemoglobin (AUC 0.474, 95% CI 0.344-0.603; p=0.694).

Conclusion: Combined anemia-hypoalbuminemia phenotypes were associated with distinct clinical and laboratory profiles but were not significantly associated with in-hospital mortality in this cohort. Neither admission albumin nor hemoglobin alone demonstrated significant discrimination for in-hospital mortality. These findings should be interpreted cautiously given the limited number of mortality events.

Keywords: Acute heart failure, anemia, hypoalbuminemia, mortality, phenotype, prognosis

Öz

%7,8 olup gruplar arasında anlamlı fark saptanmadı (Fisher'ın kesin testi, p=0,402). Yaşa göre düzeltilmiş lojistik regresyon modelinde fenotip grubu hastane içi mortalite ile anlamlı olarak ilişkili değildi (genel p=0,391) ve fenotiplerin hiçbirinde referans gruba kıyasla anlamlı ilişki saptanmadı. Bootstrap duyarlılık analizi de benzer sonuçlar verdi. Alıcı işlem karakteristiği analizinde albümin [eğri altında kalan alan (AUC) 0,594; %95 güven aralığı (GA) 0,477-0,711; p=0,157] ve hemoglobinin (AUC 0,474; %95 GA 0,344-0,603; p=0,694) hastane içi mortaliteyi ayırt etme performansı sınırlı bulundu.

Sonuç: Kombine anemi-hipoalbüminemi fenotipleri farklı klinik ve laboratuvar özellikleriyle ilişkili olmakla birlikte, bu kohortta hastane içi mortalite ile anlamlı bir ilişki göstermedi. Başvuru sırasındaki albümin ve hemoglobin düzeylerinin hiçbiri tek başına hastane içi mortalite açısından anlamlı ayırt edici performans göstermedi. Bulgular, mortalite olaylarının sınırlı sayıda olması nedeniyle dikkatle yorumlanmalıdır.

Anahtar kelimeler: Akut kalp yetersizliği, anemi, hipoalbüminemi, mortalite, fenotip, prognoz

Introduction

Acute heart failure (AHF) is among the foremost precipitants of urgent hospital admission and is burdened by substantial early mortality and high rates of return hospitalization, even in the context of evolving treatment paradigms (1). The trajectory of illness among affected individuals is highly heterogeneous, mirroring the diverse pathophysiological substrates and varying degrees of cardiac decompensation that underlie this syndrome (2). As a result, the availability of practical and dependable markers for prognostic stratification is critical to enable individualized patient management and support timely therapeutic decision-making (3).

Anemia is a prevalent comorbid condition in the heart failure population, consistently linked to impaired exercise tolerance, increased readmission frequency, and worse long-term survival (4,5). The mechanisms underlying anemia in this context are heterogeneous, encompassing contributions from renal dysfunction, iron deficiency states, chronic low-level systemic inflammation, hemodynamic fluid redistribution, and disruption of neurohumoral regulatory systems (6). In parallel, a decline in serum albumin concentration is a common biochemical accompaniment of acute cardiac decompensation, potentially reflecting malnutrition, hepatic congestion from venous hypertension, systemic inflammatory burden, and the cumulative severity of the patient's condition (7,8). A body of prior evidence has demonstrated that hypoalbuminemia confers elevated risk for adverse cardiovascular outcomes and mortality in populations with heart failure (9).

Although both abnormalities have individually demonstrated prognostic relevance, relatively little attention has been directed toward their simultaneous occurrence as an integrated clinical phenotype. Considering that anemia and hypoalbuminemia may reflect partially overlapping yet distinct aspects of systemic disease burden, evaluating their combined phenotypic patterns may provide additional insight into clinical heterogeneity and their potential relationship with short-term outcomes (10).

With this rationale, the present investigation was undertaken to evaluate the association between phenotypic combinations of anemia and hypoalbuminemia and all-cause in-hospital mortality in a cohort of patients hospitalized with AHF.

Materials and Methods

Study Design and Population

This was a single-center retrospective observational study carried out at at Mogadishu Somalia-Turkey Recep Tayyip Erdoğan Training and Research Hospital. Consecutive adult patients admitted with a confirmed diagnosis of AHF were screened for eligibility. A total of 220 patients with available clinical and laboratory data at admission formed the final study cohort.

AHF was diagnosed in accordance with established international guideline criteria, incorporating clinical

features, relevant laboratory results, and echocardiographic findings consistent with acute decompensated heart failure (1). Demographic characteristics, comorbidities, laboratory parameters, and echocardiographic data were obtained from the institutional electronic medical records.

Definitions and Study Groups

Based on anemia and hypoalbuminemia status, patients were assigned to one of four phenotypic subgroups:

1. No anemia/no hypoalbuminemia

2. Anemia only

3. Hypoalbuminemia only

4. Combined anemia and hypoalbuminemia.

Classification was based on hemoglobin and serum albumin values obtained at hospital admission. Anemia was defined according to World Health Organization thresholds as a hemoglobin level below 13 g/dL in men and below 12 g/dL in women. Hypoalbuminemia was defined as a serum albumin concentration below 3.5 g/dL.

Data Collection

The following baseline variables were systematically recorded: Age, sex, diabetes mellitus, hypertension, chronic kidney disease, left ventricular ejection fraction, hemoglobin, serum albumin, C-reactive protein (CRP), pro-B-type natriuretic peptide, and neutrophil-to-lymphocyte ratio (NLR).

Study Outcome

The primary study endpoint was all-cause mortality occurring during the index hospitalization.

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics (IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean ± standard deviation, and categorical variables as numbers and percentages. Between-group comparisons of continuous variables were performed using one-way analysis of variance. Categorical variables were compared using the Pearson chi-square test or Fisher's exact test, as appropriate according to expected cell counts.

Associations with in-hospital mortality were evaluated using binary logistic regression analysis. Univariable analyses were first performed for the anemia-hypoalbuminemia phenotype categories and selected clinically relevant variables. Because only 21 in-hospital deaths occurred, the multivariable analysis was intentionally restricted to a parsimonious model including age and the four-category anemia-hypoalbuminemia phenotype variable to reduce the risk of model overfitting. The group with neither anemia nor hypoalbuminemia served as the reference category. Results are reported as odds ratios (ORs) with 95% confidence intervals (CIs). To assess the robustness of the regression findings in the setting of the limited number of events, a sensitivity analysis using 1,000 bootstrap samples with bias-corrected and accelerated 95% CIs was also performed.

Receiver operating characteristic (ROC) curve analysis was used to evaluate the discriminatory ability of admission serum albumin and hemoglobin levels for in-hospital mortality. Death was defined as the positive outcome. Areas under the curve (AUCs) with 95% CIs were calculated. All tests were two-sided, and a p-value <0.05 was considered statistically significant.

Ethical Approval

Ethical approval was obtained from the Institutional Ethics Committee of Mogadishu Somalia-Turkey Recep Tayyip Erdoğan Training and Research Hospital (approval no: 1206, date: 02.05.2025). All procedures were conducted in accordance with the ethical principles of the Declaration of Helsinki.

Results

A total of 220 patients were included in the analysis. Based on admission hemoglobin and serum albumin levels, 39 patients (17.7%) had neither anemia nor hypoalbuminemia, 34 (15.5%) had anemia alone, 45 (20.5%) had hypoalbuminemia alone, and 102 (46.4%) had both anemia and hypoalbuminemia. Significant between-group differences were observed for hemoglobin and serum albumin levels ($p<0.001$ for both), as expected from the phenotype definitions. Baseline demographic, clinical, and laboratory characteristics according to phenotype group are presented in Table 1. Overall, 21 of 220 patients (9.5%) died during the index hospitalization. In-hospital mortality occurred in 3 of 39 patients (7.7%) with neither anemia nor hypoalbuminemia, 6 of 34 (17.6%) with anemia alone, 4 of 45 (8.9%) with hypoalbuminemia alone, and 8 of 102 (7.8%) with both abnormalities. Although mortality was numerically highest in the anemia-only group, the overall difference among phenotype groups was not statistically significant (Fisher's exact $p=0.402$). Therefore, no stepwise pattern of increasing mortality across the phenotype categories was observed.

Results of the logistic regression analyses for in-hospital mortality are presented in Table 2. In univariable analyses, none of the anemia-hypoalbuminemia phenotype categories was significantly associated with in-hospital mortality compared with patients with neither condition. Given the limited number of mortality events (n=21), the multivariable analysis was intentionally restricted to an age-adjusted parsimonious model to reduce the risk of overfitting.

In the age-adjusted model, the anemia-hypoalbuminemia phenotype variable was not significantly associated with in-hospital mortality overall (p=0.391). Compared with patients with neither abnormality, the adjusted OR was 2.431 (95% CI 0.552-10.706; p=0.240) for anemia alone, 1.211 (95% CI 0.252-5.820; p=0.811) for hypoalbuminemia alone, and 0.909 (95% CI 0.224-3.689; p=0.894) for combined anemia and hypoalbuminemia. Age was also not statistically significant in the adjusted model (OR 1.029, 95% CI 0.996-1.063; p=0.083). A sensitivity analysis based on 1,000 bootstrap samples yielded consistent findings, with no phenotype category reaching statistical significance; the wide bootstrap CIs further indicated substantial uncertainty associated with the limited number of mortality events.

ROC analysis demonstrated limited discriminatory performance of both admission serum albumin and hemoglobin for in-hospital mortality (Figure 1). The AUC was 0.594 (95% CI 0.477-0.711; p=0.157) for albumin and 0.474 (95% CI 0.344-0.603; p=0.694) for hemoglobin (n=220). Neither biomarker therefore demonstrated statistically significant discrimination for in-hospital mortality.

Table 1. Baseline characteristics according to combined anemia and hypoalbuminemia phenotypes

Variable	No anemia/no hypoalb. (n=39)	Anemia only (n=34)	Hypoalb. only (n=45)	Both (n=102)	p-value
Age, years	58.69±10.08	60.12±14.53	56.33±15.13	61.05±15.87	0.333
Male, n (%)	27 (69.2)	20 (58.8)	30 (66.7)	61 (59.8)	0.660
Diabetes mellitus, n (%)	29 (74.4)	22 (64.7)	27 (60.0)	51 (50.0)	0.054
Hypertension, n (%)	26 (66.7)	20 (58.8)	21 (46.7)	51 (50.0)	0.217
Chronic kidney disease, n (%)	6 (15.4)	7 (20.6)	3 (6.7)	13 (12.7)	0.323
Hemoglobin, g/dL	14.25±1.23	10.72±1.28	13.85±1.10	10.25±1.50	<0.001
Albumin, g/dL	4.20±0.38	4.09±0.47	2.60±0.53	2.60±0.52	<0.001
CRP, mg/L	44.79±38.71	68.09±70.12	87.16±96.66	75.66±69.31	0.054
ProBNP, pg/mL	5685±3231	6114±3468	6541±3568	5743±3457	0.572
Ejection fraction, %	38.85±22.26	37.21±18.55	43.44±21.69	44.26±19.48	0.232
NLR	3.11±2.11	7.93±18.18	5.90±6.21	7.77±13.81	0.199
Mortality, n (%)	3 (7.7)	6 (17.6)	4 (8.9)	8 (7.8)	0.402*

*: Data are presented as mean ± SD for continuous variables and n (%) for categorical variables. Group comparisons were performed using one-way analysis of variance or Pearson chi-square test, as appropriate. Fisher's exact test was used for the mortality comparison because of small expected cell counts. CRP: C-reactive protein, proBNP: Pro-B-type natriuretic peptide, NLR: Neutrophil-to-lymphocyte ratio, Hypoalb.: Hypoalbuminemia

Table 2. Logistic regression analysis for in-hospital mortality

Predictor	Univariable OR (95% CI)	p-value	Age-adjusted OR (95% CI)	p-value
Phenotype				
No anemia/no hypoalbuminemia	Reference	—	Reference	—
Anemia only	2.571 (0.590-11.198)	0.208	2.431 (0.552-10.706)	0.240
Hypoalbuminemia only	1.171 (0.245-5.585)	0.843	1.211 (0.252-5.820)	0.811
Both anemia and hypoalbuminemia	1.021 (0.257-4.065)	0.976	0.909 (0.224-3.689)	0.894
Age, per year	1.028 (0.996-1.060)	0.086	1.029 (0.996-1.063)	0.083
Ejection fraction, per 1%	0.997 (0.975-1.019)	0.796	—	—
Chronic kidney disease	2.279 (0.765-6.784)	0.139	—	—

OR: Odds ratio, CI: Confidence interval. The reference category was no anemia/no hypoalbuminemia. Because only 21 in-hospital deaths occurred, the multivariable model was intentionally restricted to age and phenotype group to reduce the risk of overfitting. Ejection fraction and chronic kidney disease were therefore evaluated only in univariable analyses. A sensitivity analysis using 1.000 bootstrap samples with bias-corrected and accelerated 95% confidence intervals did not materially alter the interpretation of the regression findings

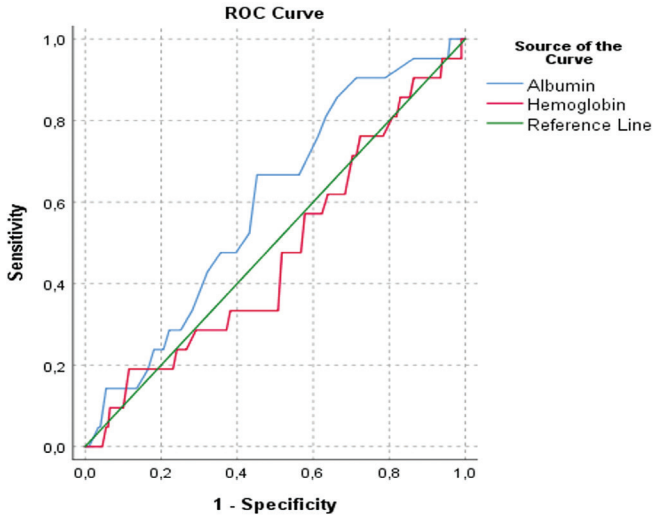

Figure 1. Receiver operating characteristic (ROC) curves for serum albumin and hemoglobin in discriminating in-hospital mortality. The areas under the curve were 0.594 (95% CI 0.477-0.711; p=0.157) for albumin and 0.474 (95% CI 0.344-0.603; p=0.694) for hemoglobin (n=220)

CI: Confidence interval

Discussion

This study evaluated the association between combined anemia-hypoalbuminemia phenotypes and in-hospital mortality among patients hospitalized with AHF. The principal finding was that, after reconstruction of the phenotype groups and reanalysis of the complete 220-patient cohort, in-hospital mortality did not differ significantly across the four phenotype categories. Although mortality was numerically highest among patients with anemia alone, none of the phenotype categories was significantly associated with mortality in either univariable or age-adjusted logistic regression analyses. In addition, neither admission serum albumin nor hemoglobin demonstrated statistically significant discriminatory ability for in-hospital mortality. These findings suggest that, in the present cohort, anemia-hypoalbuminemia phenotyping primarily characterized clinical and laboratory heterogeneity rather than providing independent short-term prognostic discrimination.

Anemia is widely recognized as a frequent and clinically meaningful accompaniment of heart failure and has been associated with reduced exercise capacity, increased rates of rehospitalization, and poorer long-term survival (4,5). Its pathophysiology in heart failure is multifactorial and may involve iron deficiency, persistent systemic inflammation, impaired renal function, neurohormonal dysregulation, and hemodilution related to volume overload (6,11,12). In the present study, patients with anemia alone had the numerically highest in-hospital mortality rate; however, this difference was not statistically significant, and the association remained non-significant after adjustment for age. The wide CIs around the effect estimates indicate substantial uncertainty and preclude firm conclusions regarding the independent short-term prognostic contribution of anemia in this cohort.

Hypoalbuminemia is also frequently encountered in patients with heart failure and may reflect several interrelated processes, including systemic inflammation, hepatic congestion secondary to elevated venous pressure, nutritional impairment, and greater overall disease burden (7,8). Previous studies have reported associations between lower serum albumin levels and adverse outcomes in both acute and chronic heart failure populations (9,13). However, the present findings did not demonstrate a significant association between the isolated hypoalbuminemia phenotype and in-hospital mortality. Importantly, this should not be interpreted as contradicting the established biological and prognostic relevance of hypoalbuminemia in broader heart failure populations. Rather, it indicates that such an association could not be demonstrated in this particular cohort, in which the number of in-hospital deaths was limited.

The combined presence of anemia and hypoalbuminemia was likewise not associated with an increased risk of in-hospital mortality compared with the reference phenotype. Although anemia and hypoalbuminemia may arise through partially shared mechanisms, including inflammation, congestion, renal dysfunction, nutritional impairment, and neurohormonal activation (14), the present data do not support an additive prognostic effect of their coexistence. At the same time, the absence of a statistically significant association should not be interpreted as evidence that anemia adds no risk in patients with hypoalbuminemia. The relatively small number of mortality events and the resulting wide CIs limit the ability to distinguish between true absence of an association and insufficient statistical precision.

Inflammatory markers also varied numerically across the phenotype groups. CRP levels tended to be higher in groups characterized by hypoalbuminemia, whereas NLR values showed considerable variability. However, the between-group differences did not reach conventional statistical significance in the reanalysis. Therefore, although systemic inflammation remains a biologically plausible link among

AHF, anemia, and hypoalbuminemia, the present findings do not establish a graded inflammatory pattern across the four phenotypes.

ROC analysis provided additional perspective on the prognostic performance of the individual biomarkers. Admission serum albumin yielded an AUC of 0.594, whereas hemoglobin yielded an AUC of 0.474; neither value was statistically significant. These findings indicate limited discriminatory performance of either biomarker alone for in-hospital mortality in this cohort. Importantly, the current analysis also does not demonstrate that the combined phenotype classification provides incremental prognostic value beyond albumin, hemoglobin, or established clinical assessment. Accordingly, the phenotype approach should be regarded as an exploratory method of characterizing clinical heterogeneity rather than as a validated risk-stratification tool.

Clinical Implications

Hemoglobin and serum albumin are routinely available laboratory measurements obtained at hospital admission and can readily identify clinically recognizable anemia-hypoalbuminemia phenotypes without additional testing or cost. Nevertheless, the present findings do not support using these phenotype categories as a stand-alone tool for short-term mortality prediction in patients with AHF. Their potential clinical relevance may instead lie in identifying different patterns of hematologic and nutritional abnormalities that warrant broader clinical assessment. Larger prospective studies with a greater number of outcome events are required to determine whether these phenotypes provide prognostic information beyond established clinical risk factors.

Study Limitations

Several limitations of this study should be acknowledged. First, its retrospective, single-center design limits the generalizability of the findings to broader patient populations and different healthcare settings. Second, although the overall cohort included 220 patients, only 21 in-hospital deaths occurred. This limited number of outcome events reduced statistical power, particularly for comparisons across the four phenotype groups, and resulted in wide CIs around several effect estimates. To reduce the risk of overfitting, the multivariable analysis was therefore intentionally restricted to a parsimonious age-adjusted model. Although a bootstrap sensitivity analysis using 1.000 samples yielded findings consistent with the primary regression analysis, the substantial uncertainty surrounding the estimates should be considered when interpreting the results. Accordingly, the absence of statistically significant associations should not be interpreted as evidence of equivalence or absence of clinically relevant effects.

Furthermore, potentially important confounding variables, including iron status parameters, formal nutritional assessment scores, and a broader range of inflammatory biomarkers, were not systematically available. Residual confounding therefore cannot be excluded. Finally, outcome assessment was restricted to all-cause mortality during the index hospitalization, and no post-discharge or long-term outcomes were evaluated. Larger, prospective, multicenter studies with a greater number of outcome events and longer follow-up are needed to determine whether anemia-hypoalbuminemia phenotypes have prognostic value beyond established clinical risk factors.

Conclusion

In patients hospitalized with AHF, combined anemia-hypoalbuminemia phenotypes were not significantly associated with in-hospital mortality. Neither isolated hypoalbuminemia nor the combined presence of anemia and hypoalbuminemia independently predicted in-hospital death, and admission serum albumin and hemoglobin demonstrated limited discriminatory performance when evaluated individually. Although these readily available laboratory parameters may help characterize hematologic and nutritional heterogeneity in AHF, the present findings do not support their use, either individually or as combined phenotypes, as stand-alone tools for short-term mortality risk stratification. Larger prospective studies with a greater number of outcome events are needed to clarify their potential prognostic value.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Institutional Ethics Committee of Mogadishu Somalia-Turkey Recep Tayyip Erdoğan Training and Research Hospital (approval no: 1206, date: 02.05.2025). All procedures were conducted in accordance with the ethical principles of the Declaration of Helsinki.

Informed Consent: It was not deemed necessary, as the study was retrospective.

Acknowledgments

OpenAI's ChatGPT was used only for language editing and manuscript organization support. All scientific content,

interpretation, and final approval were performed entirely by the authors.

Footnotes

Authorship Contributions

Surgical and Medical Practices: C.B.A., N.S.Y., Concept: C.B.A., M.Z.A., Design: C.B.A., M.Z.A., Data Collection or Processing: C.B.A., M.K.T., Analysis or Interpretation: M.Z.A., M.K.T., N.S.Y., Literature Search: M.K.T., Writing: N.S.Y.

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

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

References

1. McDonagh TA, Metra M, Adamo M, Gardner RS, Baumbach A, Böhm M, et al.; ESC Scientific Document Group. 2021 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure. Eur Heart J. 2021;42(36):3599-3726. Erratum in: Eur Heart J. 2021;42(48):4901.
2. Gheorghiade M, Pang PS. Acute heart failure syndromes. J Am Coll Cardiol. 2009;53(7):557-573.
3. Metra M, Teerlink JR. Heart failure. Lancet. 2017;390(10106):1981-1995.
4. Felker GM, Adams KF Jr, Gattis WA, O'Connor CM. Anemia as a risk factor and therapeutic target in heart failure. J Am Coll Cardiol. 2004;44(5):959-966.
5. Horwich TB, Fonarow GC, Hamilton MA, MacLellan WR, Borenstein J. Anemia is associated with worse symptoms, greater impairment in functional capacity and a significant increase in mortality in patients with advanced heart failure. J Am Coll Cardiol. 2002;39(11):1780-1786.
6. Tang WH, Van Lente F, Shrestha K, Troughton RW, Francis GS, Tong W, et al. Impact of myocardial function on cystatin C measurements in chronic systolic heart failure. J Card Fail. 2008;14(5):394-399.
7. Arques S, Ambrosi P. Human serum albumin in the clinical syndrome of heart failure. J Card Fail. 2011;17(6):451-458.
8. Bonilla-Palomas JL, Gámez-López AL, Moreno-Conde M, López-Ibáñez MC, Anguita-Sánchez MP, Gallego de la Sacristana A, et al. Hypoalbuminemia in acute heart failure patients: causes and its impact on hospital and long-term mortality. J Card Fail. 2014;20(5):350-358.
9. Ancion A, Allepaerts S, Oury C, Gori AS, Piérard LA, Lancellotti P. Serum albumin level and hospital mortality in acute non-ischemic heart failure. ESC Heart Fail. 2017;4(2):138-145.
10. Felker GM, Gattis WA, Leimberger JD, Adams KF, Cuffe MS, Gheorghiade M, et al. Usefulness of anemia as a predictor of death and rehospitalization in patients with decompensated heart failure. Am J Cardiol. 2003;92(5):625-628.
11. Westenbrink BD, Voors AA, de Boer RA, Schuringa JJ, Klinkenberg T, van Veldhuisen DJ. Bone marrow dysfunction in chronic heart failure patients. Eur J Heart Fail. 2010;12(7):676-684.
12. Uthamalingam S, Kandala J, Daley M, Patvardhan E, Capodilupo R, Moore SA, et al. Serum albumin and mortality in acutely decompensated heart failure. Am Heart J. 2010;160(6):1149-1155.
13. von Haehling S, Jankowska EA, van Veldhuisen DJ, Ponikowski P, Anker SD. Iron deficiency and cardiovascular disease. Nat Rev Cardiol. 2015;12(11):659-669.
14. Núñez J, Núñez E, Bodí V, Sanchis J, Miñana G, Mainar L, et al. Usefulness of the neutrophil to lymphocyte ratio in predicting long-term mortality in ST segment elevation myocardial infarction. Am J Cardiol. 2008;101(6):747-752.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):366-375
DOI: 10.4274/BMB.galenos.2026.62533

Investigation of Childhood Drowning Cases Presenting to the Emergency Department of a Tertiary Care Paediatric Hospital Near the Sea Coast

Deniz Kıyısına Yakın Bir Üçüncü Basamak Pediyatri Hastanesinin Acil Servisine Başvuran Çocuk Boğulma Olgularının İncelenmesi

Emre Aygün[1], Mustafa Safa Kasım[1], Mustafa Oğur[2], Ercüment Petmezci[2], Nazan Dalgıç[3]

[1]University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital, Department of Pediatrics, İstanbul, Turkey
[2]University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital, Department of Pediatrics, Division of Pediatric Intensive Care, İstanbul, Turkey
[3]University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital, Department of Pediatrics, Division of Pediatric Infectious Diseases, İstanbul, Turkey

Abstract

Objective: This study aimed to evaluate the epidemiological and clinical features of childhood drowning cases admitted to a seaside tertiary paediatric hospital emergency department and to assess the predictive value of Szpilman classification.

Method: Between 01.01.2021 and 18.09.2024 32 patients aged >1 month and <18 years presenting with drowning were included. Patients were graded according to Szpilman classification (Grade 1-6) and analysed in three groups (Grade 1-2, 3-4 and 5-6). Demographics, vital signs, laboratory and radiological findings, hospitalisation and clinical outcomes were recorded. Discharged and hospitalised patients were compared using univariable exploratory analyses.

Results: The mean age was 9.96±5.53 years and 68.8% were male. Most cases occurred at sea (81.3%), during summer (75%), and 65.6% of presentations were in the noon/afternoon hours. The most common Szpilman grades were Grade 1 (37.5%) and Grade 4 (25.0%). 62.5% of the patients were hospitalised; all Grade 1 patients were discharged and all grade ≥2 patients were hospitalised. Szpilman grade ≥4 (60.0% vs. 0%), pathological chest X-ray (85.0% vs. 0%) and higher lactate (5.79 vs. 2.69 mmol/L) were associated with hospitalisation (p≤0.001

Öz

Amaç: Bu çalışmanın amacı, deniz kenarındaki bir üçüncü basamak pediyatri hastanesinin acil servisine başvuran çocuk boğulma olgularının epidemiyolojik ve klinik özelliklerini değerlendirmek ve Szpilman sınıflandırmasının öngörü değerini belirlemektir.

Yöntem: 01.01.2021-18.09.2024 tarihleri arasında boğulma olgusuyla başvuran 1 aylıktan büyük ve 18 yaşından küçük 32 hasta çalışmaya dahil edildi. Hastalar Szpilman sınıflandırmasına göre derecelendirildi (Derece 1-6) ve üç grupta (Derece 1-2, 3-4 ve 5-6) incelendi. Demografik bilgiler, yaşamsal bulgular, laboratuvar ve radyolojik bulgular, yatış durumu ve klinik sonuçlar kaydedildi. Taburcu edilen ve yatırılan hastalar tek değişkenli keşifsel analizlerle karşılaştırıldı.

Bulgular: Hastaların yaş ortalaması 9,96±5,53 yıl idi ve hastaların %68,8'i erkekti. Olguların çoğu denizde (%81,3) ve yaz aylarında (%75) meydana geldi; başvuruların %65,6'sı öğle/öğleden sonra saatlerindeydi. En sık görülen Szpilman dereceleri Derece 1 (%37,5) ve Derece 4 (%25,0) idi. Hastaların %62,5'i yatırıldı; tüm 1. derece hastalar taburcu edilirken, tüm 2. derece ve üzeri hastalar yatırıldı. Szpilman derecesi ≥4 (%60,0'a karşı %0), patolojik akciğer grafisi

Address for Correspondence: Emre Aygün, University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital, Department of Pediatrics, İstanbul, Turkey
E-mail: emreaygun88@gmail.com **ORCID:** orcid.org/0000-0002-0546-327X
Received: 21.05.2026 **Accepted:** 15.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Aygün E, Kasım MS, Oğur M, Petmezci E, Dalgıç N. Investigation of childhood drowning cases presenting to the emergency department of a tertiary care paediatric hospital near the sea coast. Bagcilar Med Bull. 2026;11(3):366-375

Abstract

for all). Hospital stay increased with Szpilman grade (Grade 1-2: 2.00 days, Grade 3-4: 4.46 days, Grade 5-6: 39.50 days; p<0.001). Poor outcome (death or sequelae) occurred in four patients (12.5%), three of whom were Grade 5-6 (p=0.003).

Conclusion: The Szpilman classification is a useful tool for estimating clinical outcomes in childhood drowning. Higher Szpilman grade, elevated lactate and pathological radiological findings were associated with hospitalisation. Early recognition of factors associated with poor prognosis may contribute to timely intervention.

Keywords: Drowning, emergency service, pediatrics, prognosis, Szpilman classification

Öz

(%85,0'a karşı %0) ve yüksek laktat (5,79'a karşı 2,69 mmol/L) yatış ile ilişkiliydi (tümü için p≤0,001). Hastanede kalış süresi Szpilman derecesiyle arttı (1-2. derece: 2,00 gün, 3-4. derece: 4,46 gün, 5-6. derece: 39,50 gün; p<0,001). Kötü sonuç (ölüm veya sekel) dört hastada (%12,5) görüldü ve bunların üçü Derece 5-6 idi (p=0,003).

Sonuç: Szpilman sınıflandırması, çocukluk çağı boğulma olgularında klinik sonuçları öngörmede yararlı bir araçtır. Yüksek Szpilman derecesi, yüksek laktat düzeyi ve patolojik radyolojik bulgular yatış ile ilişkiliydi. Kötü prognozla ilişkili faktörlerin erken tanınması, zamanında müdahaleye katkıda bulunabilir.

Anahtar kelimeler: Acil servis, boğulma, pediyatri, prognoz, Szpilman sınıflandırması

Introduction

Worldwide, drowning is an important cause of preventable death in childhood. According to World Health Organization data, approximately 372,000 people die annually due to drowning, and the actual number is thought to be higher due to lack of data in low- and middle-income countries (1). It is reported as the second most common cause of death in children in some countries (2). In children under 15 years of age, drowning causes 140,219 deaths per year, ranking third after AIDS and meningitis (3). In the paediatric group, children aged 1-4 years and adolescents (15-19 years) constitute the highest risk groups (4). Cases occur during daily life in low- and middle-income countries and during recreational activities in high-income countries (1). Drowning is defined as "the process of respiratory impairment caused by immersion in liquid or obstruction of the airways by liquid" (5). Hypoxia and hypothermia can lead to significant complications (6). In children, body structure characteristics cause rapid heat loss and breathing difficulties (7). Hypoxia, hypercapnia and acidosis that develop after drowning can lead to long-term neurological damage by affecting brain functions (6). Cases are more common in boys, in summer and during water activities (2).

Although drowning is a global public health problem, data regarding its epidemiology, clinical course, prognostic factors, and optimal management in children remain limited. Few studies have evaluated pediatric drowning using severity assessment tools such as the Szpilman classification (8). Furthermore, evidence regarding the safe discharge of asymptomatic or mildly symptomatic patients, the role of non-invasive ventilation (NIV) strategies, and predictors of hospitalization and respiratory support requirements remains insufficient (8). Although factors such as submersion duration and water temperature are known to influence outcomes, these variables are frequently unavailable in clinical studies, and the impact of hypothermia and age-specific differences on prognosis remains unclear (8). In addition, prognostic parameters have not been adequately investigated in pediatric drowning victims, and no standardized prognostic model has yet been established (6). Therefore, further studies are needed to improve risk stratification and clinical decision-making in childhood drowning cases.

The hypothesis of our study was that epidemiological and clinical factors would be strongly associated with disease severity and prognosis in children admitted to our seaside tertiary paediatric hospital for drowning. Our main aim was to retrospectively evaluate the epidemiological characteristics, clinical features, management strategies, and outcomes of paediatric drowning cases, as well as to assess the predictive value of the Szpilman classification and identify factors associated with disease severity and prognosis.

Materials and Methods

Study Population and Sample

This study was designed as a single-centre, retrospective study conducted at a tertiary care paediatric hospital. Between 01.01.2021 and 18.09.2024, the electronic medical records of 32 patients admitted to the paediatric emergency department with the complaint of drowning were evaluated. All patients who were admitted to the hospital within the specified date range and met the study criteria were included in the study without sample size calculation. Inclusion criteria were age >1 month and <18 years and presentation to the paediatric emergency department with drowning. Patients aged ≤1 month or ≥18 years were excluded. Drowning is defined as impairment of respiratory function as a result of contact of the respiratory

tract with water in any water environment (sea, pool, lake, etc.). Szpilman classification was used for severity assessment. Patients were classified as Grade 1 (normal pulmonary auscultation with coughing), Grade 2 (abnormal pulmonary auscultation with rales in some pulmonary fields), Grade 3 (acute pulmonary oedema without arterial hypotension), Grade 4 (acute pulmonary oedema with arterial hypotension), Grade 5 (isolated respiratory arrest/apnoea) and Grade 6 (cardiopulmonary arrest). The Szpilman classification was assigned retrospectively by the investigators after review of the medical records and was not used for clinical decision-making during patient management.

Operating Procedures

Data collection was performed retrospectively using the relevant International Classification of Diseases, 10th Revision (ICD-10) diagnosis codes (T75.1: Drowning and non-fatal submersion) from the hospital electronic medical record system (health information management system). Demographic characteristics (age, gender), clinical characteristics (nature of the water, season, month, day, hour, presence of chronic disease, initial cardiopulmonary resuscitation (CPR) application, Szpilman classification), vital signs [peak heart rate, respiratory rate, body temperature, oxygen saturation, Glasgow Coma scale (GCS)], laboratory values (arterial blood gas parameters, blood biochemistry) and radiological findings were recorded using a standardised data collection form. Laboratory analyses were performed in the central laboratory of our hospital in accordance with standard protocols. Radiometer ABL800 FLEX analyser was used for arterial blood gas analyses and Beckman Coulter AU5800 autoanalyzer was used for biochemical parameters. Patients were followed up until discharge from the hospital or death. Clinical outcome was recorded as full recovery, recovery with sequelae or death, and poor outcome was defined as death or recovery with sequelae. Radiological evaluation was performed by radiology specialists of our hospital in accordance with standard protocols, and findings such as infiltration, oedema and atelectasis on chest radiographs were classified as "pathological findings". The reliability of the data collection tools was tested with a pilot study and compared by data extraction by two independent researchers. Owing to the retrospective design, some vital signs and laboratory values were not available for every patient, and analyses were performed with the available data.Troponin was measured in 22 patients (19 hospitalised and 3 discharged). Oxygen saturation could not be measured in patients with cardiopulmonary arrest, and these patients were excluded from the analysis of this variable.

Response Protocol

Patients were divided into two groups according to their clinical status: those discharged from the emergency department and those hospitalised. Randomization was not performed. Hospitalisation and pediatric intensive care unit (PICU) admission decisions were made by the treating physicians based on the overall clinical assessment, including respiratory status, oxygen saturation, neurological findings, radiological abnormalities, and the anticipated risk of clinical deterioration.

The protocols applied to the patients admitted to the intensive care unit were performed in accordance with the standard protocols of the paediatric intensive care unit of our hospital. According to the need for respiratory support, patients were divided into room air, NIV [high-flow nasal cannula (HFNC) or bilevel positive airway pressure (BiPAP)] and invasive mechanical ventilation groups. Antibiotic treatment was initiated when drowning-associated pneumonia was clinically suspected based on respiratory symptoms, abnormal chest radiographic findings, fever, laboratory markers of infection, and physician assessment. Sulbactam-ampicillin, piperacillin-tazobactam, and meropenem were the most commonly prescribed antibiotics. The duration of intensive care unit follow-up and discharge decision were based on the criteria of improvement of the patient's clinical condition, stabilisation of vital findings and normalisation of laboratory parameters.

Statistical Analysis

IBM SPSS Statistics 25.0 software was used in statistical analyses. Descriptive statistics were presented as number-percentage for categorical variables and mean-standard deviation for continuous variables; the median was also given for length of hospital stay. Demographic and clinical characteristics were presented descriptively, without hypothesis testing. Normal distribution was evaluated by Kolmogorov-Smirnov test. Hospitalised patients and patients discharged from the emergency department were compared using the independent samples t-test or the Mann-Whitney U test for continuous variables, according to the distribution of the data, and Fisher's exact test for categorical variables. Because the number of patients in several individual Szpilman grades was very small, patients were grouped into three clinically meaningful categories

(Grade 1-2, Grade 3-4 and Grade 5-6). Continuous variables were compared among these groups using One-Way ANOVA and categorical variables using the Fisher-Freeman-Halton exact test. Given the small sample size, multivariable logistic regression analysis was not performed. Factors associated with hospitalisation were examined as exploratory univariable associations using Fisher's exact test, and odds ratios were not estimated because of zero cells. Since only four patients had a poor outcome, this group was compared with patients with good outcome descriptively and with Fisher's exact test only. A p-value <0.05 was considered statistically significant.

Ethical Considerations

Ethical approval for this study was obtained from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital (approval number: 2024/238, date: October 15, 2024). Due to the retrospective nature of the study using anonymized medical records, the requirement for informed consent was waived by the ethics committee. Personal data were anonymised and coded and analysed to protect the privacy of patients. The study was conducted in accordance with the principles of the Declaration of Helsinki and the good clinical practice guidelines.

Results

The mean age of the 32 patients was 9.96±5.53 years and the majority were boys (68.8%). Most of the cases were drowning in the sea (81.3%), while drowning in the pool was less common (18.8%). Seasonally, the incidents were concentrated in the summer months (75%) and occurred most frequently in July (40.6%). Most of the admissions took place during lunch and afternoon hours (65.6%). Most patients had no chronic diseases (87.5%) and initial CPR was required in only 15.6% of cases. According to the Szpilman classification, most patients were Grade 1 (37.5%), followed by Grade 4 (25.0%) and Grade 3 (15.6%) cases (Table 1).

Compared with patients discharged from the emergency department, hospitalised patients had lower oxygen saturation (86.83% vs. 98.27%, p=0.001), lower GCS scores (12.14 vs. 15.00, p=0.007), lower body temperature (36.28 °C vs. 36.69 °C, p=0.015) and a higher respiratory rate (37.00 vs. 21.45/min, p<0.001). Arterial blood gas parameters showed more pronounced acidosis (pH 7.18 vs. 7.38, p=0.003), higher lactate (5.79 vs. 2.69 mmol/L, p=0.001) and lower bicarbonate (16.85 vs. 22.18 mmol/L, p=0.001) values in hospitalised patients. Glucose (153.55 vs. 101.73 mg/dL,

Table 1. Demographic and clinical characteristics of the patients (n=32)

Parameter	n (%)
Age, years, mean ± SD (min-max)	9.96±5.53 (2.00-17.25)
Sex	
Male	22 (68.8)
Female	10 (31.3)
Type of water	
Sea	26 (81.3)
Pool	6 (18.8)
Season	
Spring (May)	1 (3.1)
Summer (June-August)	24 (75.0)
Autumn (September-October)	6 (18.8)
Winter (December)	1 (3.1)
Month	
May	1 (3.1)
June	4 (12.5)
July	13 (40.6)
August	7 (21.9)
September	4 (12.5)
October	2 (6.3)
December	1 (3.1)
Day of the week	
Monday	5 (15.6)
Tuesday	4 (12.5)
Wednesday	7 (21.9)
Thursday	2 (6.3)
Friday	1 (3.1)
Saturday	7 (21.9)
Sunday	6 (18.8)
Time of presentation	
Morning	4 (12.5)
Noon/afternoon	21 (65.6)
Evening	7 (21.9)
Chronic disease	4 (12.5)
Initial CPR	5 (15.6)
Szpilman grade	
Grade 1	12 (37.5)
Grade 2	3 (9.4)
Grade 3	5 (15.6)
Grade 4	8 (25.0)
Grade 5	1 (3.1)
Grade 6	3 (9.4)

Data are presented as n (%) unless otherwise stated. CPR: Cardiopulmonary resuscitation, SD: Standard deviation

p=0.014), aspartate aminotransferase (89.60 vs. 30.82 U/L, p=0.013) and alanine aminotransferase (48.85 vs. 15.18 U/L, p=0.010) levels were also higher in hospitalised patients, whereas peak heart rate, sodium, urea, creatinine and troponin did not differ significantly between the groups. Radiological findings were pathological in 85% of the hospitalised patients, whereas the findings were normal in all discharged patients (p<0.001) (Table 2).

62.5% of the patients were hospitalised and all were followed up in the PICU. The mean duration of hospitalisation was 11.10±19.28 days (median: 4.00 days). When the distribution of respiratory support was analysed, 37.5% of the patients were monitored with room air, 31.3% with HFNC, 15.6% with BiPAP and 15.6% with invasive mechanical ventilation support. All hospitalised patients received respiratory support [NIV: 75%, invasive ventilation (IV): 25%], while those discharged were followed up on room air. Intubation was required in a total of 5 patients (15.6%). Antibiotic use was considered necessary in half of the patients (50%) and sulbactam-ampicillin (68.8%) was preferred most frequently (Table 3).

Szpilman grades were analysed in three groups because of the small number of patients in individual grades. All Grade 1 patients were discharged from the emergency department, whereas all Grade 2 and above patients were hospitalised; accordingly, the hospitalisation rate was 20.0% in the Grade 1-2 group and 100% in the Grade 3-4 and Grade 5-6 groups (p<0.001). The mean length of hospital stay was 2.00±0.00 days in Grade 1-2, 4.46±1.28 days in Grade 3-4 and 39.50±31.61 days in Grade 5-6 (p<0.001). In the Grade 1-2 group, 80.0% of the patients were followed on room air and 20.0% received HFNC. In the Grade 3-4 group, 53.8% received HFNC, 38.5% BiPAP and 7.7% invasive mechanical ventilation, while all patients in the Grade 5-6 group required invasive mechanical ventilation (p<0.001). With increasing severity, pH decreased (7.36, 7.26 and 6.79) and lactate increased (2.36, 4.29 and 13.74 mmol/L) (p<0.001 for both). Glucose rose from 102.65 to 203.50 mg/dL (p=0.001), and the GCS score fell from 14.80 to 4.00 (p<0.001). Oxygen saturation was 97.10% in Grade 1-2 and 87.60% in Grade 3-4; in the Grade 5-6 group, it could be measured in only one patient (60%), since it was not measurable in the three

Table 2. Vital signs and laboratory values at admission

Parameter	All patients (n=32)	Hospitalised (n=20)	Discharged from the ED (n=12)	p
Vital signs				
Peak heart rate (beats/min)	103.72±21.65 (n=29)	107.39±22.07 (n=18)	97.73±20.51 (n=11)	0.244
Respiratory rate (/min)	31.10±10.60 (n=29)	37.00±8.66 (n=18)	21.45±4.74 (n=11)	<0.001
Body temperature (°C)	36.42±0.56 (n=28)	36.28±0.63 (n=19)	36.69±0.19 (n=9)	0.015
SpO_2 (%)[a]	91.17±8.59 (n=29)	86.83±8.21 (n=18)	98.27±1.56 (n=11)	0.001
Glasgow Coma scale	13.13±3.69 (n=29)	12.14±4.26 (n=19)	15.00±0.00 (n=10)	0.007
Arterial blood gas				
pH	7.25±0.20 (n=31)	7.18±0.22 (n=20)	7.38±0.05 (n=11)	0.003
Lactate (mmol/L)	4.69±3.85 (n=31)	5.79±4.34 (n=20)	2.69±1.32 (n=11)	0.001
Bicarbonate (mmol/L)	18.74±5.36 (n=31)	16.85±5.77 (n=20)	22.18±1.66 (n=11)	0.001
Blood biochemistry				
Glucose (mg/dL)	135.16±59.23 (n=31)	153.55±64.98 (n=20)	101.73±24.45 (n=11)	0.014
Sodium (mmol/L)	139.13±4.15 (n=31)	139.75±4.10 (n=20)	138.00±4.17 (n=11)	0.621
Urea (mg/dL)	23.39±6.97 (n=31)	23.70±7.02 (n=20)	22.82±7.17 (n=11)	0.745
Creatinine (mg/dL)	0.58±0.31 (n=31)	0.57±0.26 (n=20)	0.60±0.39 (n=11)	0.822
AST (U/L)	68.74±127.17 (n=31)	89.60±155.38 (n=20)	30.82±13.96 (n=11)	0.013
ALT (U/L)	36.90±63.89 (n=31)	48.85±77.51 (n=20)	15.18±5.40 (n=11)	0.010
Troponin (ng/mL)	34.72±66.62 (n=22)	38.80±71.05 (n=19)	8.83±3.40 (n=3)	0.109
Chest X-ray				<0.001
Normal	15 (46.9)	3 (15.0)	12 (100.0)	
Pathological	17 (53.1)	17 (85.0)	0 (0.0)	

Data are presented as mean ± standard deviation or n (%). Owing to missing data, the number of patients with available measurements varied between variables and is given in parentheses for each continuous variable. Continuous variables were compared using the independent samples t-test or Mann-Whitney U test, and categorical variables using Fisher's exact test. [a]: Patients with cardiopulmonary arrest, in whom SpO_2 was not measurable, were excluded. ED: Emergency department, SpO_2: Peripheral oxygen saturation, AST: Aspartate aminotransferase, ALT: Alanine aminotransferase

Table 3. Hospitalisation and treatment characteristics

Parameter	n (%)
Emergency department outcome	
Discharged from the emergency department	12 (37.5)
Hospitalised	20 (62.5)
Unit of hospitalisation (n=20)	
Paediatric intensive care unit	20 (100.0)
Length of hospital stay, days (n=20)	
Mean ± SD	11.10±19.28
Median	4.00
Respiratory support (n=32)	
Room air	12 (37.5)
High-flow nasal cannula	10 (31.3)
Bilevel positive airway pressure	5 (15.6)
Invasive mechanical ventilation	5 (15.6)
Respiratory support in hospitalised patients (n=20)	
Non-invasive ventilation	15 (75.0)
Invasive mechanical ventilation	5 (25.0)
Antibiotic therapy	16 (50.0)
Antibiotic type (n=16)	
Sulbactam-ampicillin	11 (68.8)
Piperacillin-tazobactam	4 (25.0)
Meropenem	1 (6.3)

Data are presented as n (%) unless otherwise stated. SD: Standard deviation

patients with cardiopulmonary arrest. Chest X-ray was normal in all Grade 1-2 patients and pathological in all patients in the Grade 3-4 and Grade 5-6 groups (p<0.001). All patients in the Grade 1-2 group recovered fully. In the Grade 3-4 group, 92.3% recovered fully and 7.7% recovered with sequelae. In the Grade 5-6 group, one patient (25.0%) recovered fully, two (50.0%) recovered with sequelae and one (25.0%) died (p=0.002) (Table 4).

Factors associated with hospitalisation were evaluated as exploratory univariable associations. Szpilman grade ≥4 was present in 60.0% of hospitalised patients and in none of the discharged patients (p<0.001). Pathological chest X-ray findings were observed in 85.0% of hospitalised patients and in none of the discharged patients (p<0.001). Initial CPR was performed in 25.0% of hospitalised patients and in none of the discharged patients, but this difference was not statistically significant (p=0.130). Odds ratios could not be estimated because of zero cells (Table 5).

A poor outcome was observed in four patients (12.5%); three recovered with sequelae and one died. One of these patients was classified as Grade 4 and three as Grade 6. Szpilman grade 5-6 was more frequent in patients with poor outcome than in those with good outcome (75.0% vs. 3.6%, p=0.003), and all patients in the Szpilman Grade 5-6 group had a GCS score below 8. Pathological chest X-ray findings (100%

Table 4. Clinical characteristics and outcomes according to Szpilman grade groups

Parameter	Grade 1-2 (n=15)	Grade 3-4 (n=13)	Grade 5-6 (n=4)	p
Hospitalisation, n (%)	3 (20.0)	13 (100.0)	4 (100.0)	<0.001
Length of hospital stay, days[a]	2.00±0.00	4.46±1.28	39.50±31.61	<0.001
Respiratory support, n (%)				<0.001
Room air	12 (80.0)	0 (0.0)	0 (0.0)	
High-flow nasal cannula	3 (20.0)	7 (53.8)	0 (0.0)	
Bilevel positive airway pressure	0 (0.0)	5 (38.5)	0 (0.0)	
Invasive mechanical ventilation	0 (0.0)	1 (7.7)	4 (100.0)	
pH	7.36±0.05	7.26±0.10	6.79±0.09	<0.001
Lactate (mmol/L)	2.36±1.20	4.29±1.84	13.74±1.57	<0.001
Glucose (mg/dL)	102.65±23.41	135.02±54.09	203.50±67.90	0.001
Glasgow Coma scale	14.80±0.47	14.12±1.13	4.00±2.00	<0.001
SpO_2 (%)	97.10±2.78	87.60±3.42	60.00[b]	<0.001
Pathological chest X-ray, n (%)	0 (0.0)	13 (100.0)	4 (100.0)	<0.001
Outcome, n (%)				0.002
Full recovery	15 (100.0)	12 (92.3)	1 (25.0)	
Recovery with sequelae	0 (0.0)	1 (7.7)	2 (50.0)	
Death	0 (0.0)	0 (0.0)	1 (25.0)	

Data are presented as mean ± standard deviation or n (%). Continuous variables were compared using One-Way ANOVA and categorical variables using the Fisher-Freeman-Halton exact test. [a]: Hospitalised patients only (n=3, n=13 and n=4, respectively), [b]: Measured in one patient; not measurable in three patients with cardiopulmonary arrest. SpO_2: Peripheral oxygen saturation. Owing to missing data, the number of patients with available measurements varied between variables

vs. 46.4%, p=0.104) and hospitalisation (100% vs. 57.1%, p=0.271) did not differ significantly between the outcome groups. Owing to the very small number of patients with poor outcome, regression analysis was not performed and the findings are presented descriptively (Table 6).

Table 5. Factors associated with hospitalisation (exploratory univariable analysis)

Factor	Hospitalised (n=20)	Discharged (n=12)	p
Szpilman grade ≥4	12 (60.0)	0 (0.0)	<0.001
Pathological chest X-ray	17 (85.0)	0 (0.0)	<0.001
Initial CPR	5 (25.0)	0 (0.0)	0.130

Data are presented as n (%). P-values were obtained using Fisher's exact test. Because of the small sample size and the presence of zero cells, odds ratios were not estimated and the associations are presented as exploratory. Comparisons of continuous clinical and laboratory variables between the groups are presented in Table 2. CPR: Cardiopulmonary resuscitation

Table 6. Characteristics of patients with poor and good outcome

Factor	Poor outcome (n=4)	Good outcome (n=28)	p
Szpilman grade 5-6[a]	3 (75.0)	1 (3.6)	0.003
Pathological chest X-ray	4 (100.0)	13 (46.4)	0.104
Hospitalisation	4 (100.0)	16 (57.1)	0.271

Poor outcome was defined as death or recovery with sequelae. Data are presented as n (%). P-values were obtained using Fisher's exact test. Owing to the very small number of patients with poor outcome, no regression analysis was performed and the results are presented descriptively. [a]: All patients with Szpilman grade 5-6 had a Glasgow Coma scale score <8

Discussion

In this study, we evaluated the epidemiological characteristics, clinical presentation, management and outcomes of paediatric drowning cases admitted to a tertiary paediatric hospital located in a coastal region. Our findings support the clinical utility of the Szpilman classification as a practical tool for risk stratification, since higher grades were associated with a greater need for hospitalisation and respiratory support, longer hospital stay and more frequent poor outcome. These findings are particularly relevant given the limited literature evaluating the Szpilman classification in paediatric drowning cases (8). Because demographic, clinical, laboratory, radiological and outcome data were evaluated together in the same cohort, the classification could be compared with parameters that are readily available in the emergency department, such as lactate and chest radiography, both of which were associated with the need for hospitalisation.

The mean age of drowning cases in our study (9.96±5.53 years) was higher than the mean ages of 4.0±3.2 years reported by Loux et al. (9) and 3.9 years reported by Alam et al. (10). This difference may be related to the characteristics of our study area, where most cases occurred at sea rather than in domestic settings. Male predominance in gender distribution (68.8%) was consistent with the literature. Denny et al. (11) reported that approximately 75% of childhood drowning cases were in males, and Tyr et al. (12) reported a similarly high proportion of males (72.3% 0-4 years, 76.0% 5-9 years). Koon et al. (13) also reported male sex as the most frequently identified risk factor.

In 81.3% of our cases, drowning occurred at sea, while pool drownings accounted for 18.8%. Denny et al. (11) reported a high rate of drowning of older children in natural bodies of water (69%), while Loux et al. (9) reported that pool drownings occurred mostly in private pools (65.6%). In the review by Koon et al. (13), drownings in natural water areas were more common, and sea drownings were reported to be 85%, 90% and 76% in Iran, France and South Africa, respectively. Seasonally, cases were concentrated in summer (75%) and especially in July (40.6%). This finding was consistent with the 70% reported by Denny et al. (11) from May to August and the 82% reported by Alam et al. (10) from June to October. Most presentations (65.6%) were during the midday/afternoon hours, which was similar to the 75% observed by Denny et al. (11) between noon and 21:00. Peixoto-Pino et al. (14) reported that drownings are common in summer and in beach and pool environments, and Kamstra et al. (15) indirectly confirmed this seasonal risk with their study in the summer holiday period in Australia. In coastal regions, preventive measures should therefore focus on open water and on the summer months.

In our cohort, Grade 1 was the most common Szpilman grade (37.5%), and the overall distribution was different from the rates reported by Büyük and Çamcı (16) (Grade 1: 26.7%, Grade 2: 13.3%, Grade 5: 53.3%, Grade 6: 6.7%). Peixoto-Pino et al. (14) reported that Grade 1 cases were limited to coughing, while Grade 6 cases required CPR. All Grade 1 patients were discharged, whereas all patients with Grade 2 or higher were hospitalised. Since the Szpilman grade was assigned retrospectively and was not used in admission decisions, this pattern most likely reflects the severity that the treating physicians recognised at presentation, such as abnormal respiratory findings, hypoxaemia and altered consciousness. In other words, the classification summarised the same clinical information on which decisions were based in practice. Büyük and

Çamcı (16) reported that those discharged had a mild clinical picture (GCS 10-15) and those hospitalised had more severe conditions (GCS 3-5). Szpilman and Morgan (17) recommend that patients with grade 3-6 should be hospitalised in the intensive care unit; in our study, all hospitalised patients, including those with Grade 2, were followed in the PICU.

Length of hospital stay increased across the Szpilman groups, and the longest stays were observed in the Grade 5-6 group. Bellini et al. (18) reported that those with pathological chest X-ray findings were hospitalised longer. Poor outcome occurred in four patients, three of whom had presented with cardiopulmonary arrest (Grade 6). All patients in the Grade 5-6 group had a GCS score below 8, and this group included three of the four patients with poor outcome. Büyük and Çamcı (16) reported 83.3% mortality in the Szpilman 3-6 group. Berger et al. (5) reported that the rate of poor outcome was 70% in patients with GCS 3, 0% in patients with GCS 4-15, and the rate of poor outcome reached 100% in patients with pH <7. Reizine et al. (19) showed a strong association of the presence of cardiac arrest at baseline with 28-day mortality [adjusted hazard ratio (aHR) 11.5, 95% CI 2.51-52.43]. Our observations agree with Szpilman and Morgan's (17) emphasis on GCS-based prognostic assessment; however, with only four events, they can only be interpreted descriptively.

In our study, the distribution of respiratory support was room air 37.5%, HFNC 31.3%, BiPAP 15.6% and IV 15.6%. Şık et al. (20) reported that all 25 children with Grade 3-4 drowning-related pulmonary edema were treated with BiPAP, with none subsequently requiring invasive mechanical ventilation. The type of respiratory support changed with Szpilman grade. Most patients in the Grade 1-2 group remained on room air, HFNC and BiPAP were mainly used in the Grade 3-4 group, and all Grade 5-6 patients required IV. Non-invasive support was therefore sufficient for most patients up to Grade 4. Capela et al. (21) reported that IV was required in 83% of severe cases, whereas supplemental oxygen alone (26%) was sufficient in mild cases. Intubation was needed in 15.6% of our patients; one of them was in Grade 4 and the remaining four were in Grade 5-6. Reizine et al. (19) reported the association of first-day invasive mechanical ventilation with mortality. Oxygen saturation decreased with increasing severity and could not be measured in patients with cardiopulmonary arrest. Dezfulian et al. (22) emphasised that cardiac arrest after drowning is usually caused by severe hypoxaemia and stressed the importance of early oxygen therapy.

Hospitalised patients had lower pH and higher lactate levels than discharged patients, and both parameters worsened across the Szpilman groups, with the most pronounced acidosis and hyperlactataemia in the Grade 5-6 group. Berger et al. (5) reported that the risk of neurological damage/death was significantly higher in children with pH <7 [relative risk (RR) 12.3, 95% CI 4.17-36.49], and that the risk of poor outcome increased 9-fold in patients with lactate >14 mmol/L. Jeswani et al. (23) also showed that pH <7 was associated with poor outcomes in children. Sacco and Aquila (24) emphasised that lactate levels are significantly increased in hypoxia and indicate the severity of the condition. Since blood gas results are available within minutes of arrival, pH and lactate may complement the Szpilman grade in the early assessment of severity. Because of the small number of poor outcomes, we did not evaluate prognostic cut-off values for these parameters.

Chest radiographs were normal in all Grade 1-2 patients and pathological in all patients with Grade 3 or higher. Pathological findings were present in 85% of hospitalised patients and in none of the discharged patients, but they were not significantly associated with poor outcome. Bellini et al. (18) showed that pathological chest X-ray increased the likelihood of hospitalisation (p=0.023). Peri et al. (25) reported that pathological chest X-rays (bilateral opacities) were seen in 61% of intensive care unit patients, and Jeswani et al. (23) reported that abnormal findings were associated with poor prognosis. The lack of such an association in our cohort is probably related to the small number of patients with poor outcome. Szpilman and Morgan (17) also recommended chest X-ray to evaluate pulmonary oedema in all patients and our findings support this approach.

Antibiotics were administered to 50% of the patients and sulbactam-ampicillin (68.8%) was preferred most frequently. Although treatment was started only when drowning-associated pneumonia was suspected, this rate is relatively high and may reflect a low threshold for treatment in patients with abnormal chest radiographs. Cousin and Pittet (26) reported that amoxicillin-clavulanate (78-94%) was mostly used in drowning-related pneumonia.

Study Limitations

Limitations of our study include its retrospective single-centre design and relatively small sample size, which may limit the generalisability of the findings. The low number of patients with poor outcome (n=4) and the small number of patients in some Szpilman grades did not allow

multivariable analysis. The associations reported here are therefore exploratory and should not be interpreted as independent predictors of hospitalisation or outcome. Another limitation is that the Szpilman classification was assigned retrospectively based on documented clinical findings and was not routinely recorded during patient management. Some vital signs and laboratory values were missing because of the retrospective design, and oxygen saturation could not be measured in patients with cardiopulmonary arrest. In addition, important variables known to influence prognosis, such as submersion duration, water temperature, bystander resuscitation quality and prehospital interventions, were not consistently available in the medical records and therefore could not be analysed. Patients were followed only until hospital discharge, so long-term neurological outcomes could not be assessed. Despite these limitations, the study provides detailed clinical, laboratory, radiological and outcome data from a paediatric cohort and contributes additional evidence regarding the prognostic value of the Szpilman classification in childhood drowning. Future multicentre prospective studies with larger sample sizes and more comprehensive data collection are needed to validate these findings and support the development of evidence-based guidelines for paediatric drowning management.

Conclusion

We found that the Szpilman classification was useful for estimating the clinical course and outcome of drowning children treated in a coastal paediatric referral hospital. Drowning cases were more common in boys, in the summer months and at sea. Higher Szpilman grade, elevated lactate levels and pathological chest radiographic findings were associated with the need for hospitalisation, and poor outcome was seen mainly in patients presenting with cardiopulmonary arrest. As the number of patients was small, these associations should be confirmed in larger prospective studies. Early recognition of patients at risk of poor outcome may contribute to timely intervention and the prevention of possible complications.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital (approval number: 2024/238, date: October 15, 2024).

Informed Consent: Due to the retrospective nature of the study using anonymized medical records, the requirement for informed consent was waived by the ethics committee.

Footnotes

Authorship Contributions

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

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

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

References

1. Peden AE, Franklin RC, Leggat PA. Fatal river drowning: the identification of research gaps through a systematic literature review. Inj Prev. 2016;22(3):202-209.
2. Davey M, Callinan S, Nertney L. Identifying risk factors associated with fatal drowning accidents in the paediatric population: a review of international evidence. Cureus. 2019;11(11):e6201.
3. Niamsanit S, Uppala R, Sitthikarnkha P, Techasatian L, Saengnipanthkul S, Thepsuthammarat K, et al. The epidemiology and outcomes of hospitalised drowning in Thai children: a national data analysis 2015-2019. Scand J Trauma Resusc Emerg Med. 2024;32:98.
4. Pellegrino F, Raffaldi I, Rossi R, De Vito B, Pagano M, Garelli D, et al. Epidemiology, clinical aspects, and management of paediatric drowning. Ital J Pediatr. 2023;49:74.
5. Berger S, Siekmeyer M, Petzold-Quinque S, Kiess W, Merkenschlager A. Drowning and nonfatal drowning in children and adolescents: a subsequent retrospective data analysis. Children (Basel). 2024;11(4):439.
6. Raess L, Darms A, Meyer-Heim A. Drowning in children: retrospective analysis of incident characteristics, predicting parameters, and long-term outcome. Children (Basel). 2020;7(7):70.
7. Batista LL, Takita SY, Oliveira CMS, Martin JG, Fioretto JR, Bonatto RC, et al. Drowning in pediatrics: literature review. Resid Pediatr. 2023;13(1):1-10.
8. Şık N, Şenol HB, Öztürk A, Yılmaz D, Duman M. A reappraisal of childhood drowning in a paediatric emergency department. Am J Emerg Med. 2021;41:90-95.
9. Loux T, Mansuri F, Brooks SE, Slye N, Lewis B, Lu Y, et al. Factors associated with pediatric drowning admissions and outcomes at a trauma center, 2010-2017. Am J Emerg Med. 2021;39:86-91.
10. Alam E, Al Hattawi KS, Akter H, Alam J, Alvarez E, Sufi F, et al. Socioeconomic, demographic and environmental factors of child drownings in Northern Bangladesh. Inj Prev. 2025:1-10. Epub ahead of print.

11. Denny SA, Quan L, Gilchrist J, McCallin T, Shenoi R, Yusuf S, et al. Prevention of drowning. Pediatrics. 2021;148(2):e2021052227.

12. Tyr A, Molander E, Bäckström B, Claesson A, Zilg B. Unintentional drowning fatalities in Sweden between 2002 and 2021. BMC Public Health. 2024;24:3185.

13. Koon W, Peden A, Lawes JC, Brander RW. Coastal drowning: a scoping review of burden, risk factors, and prevention strategies. PLoS One. 2021;16(2):e0246034.

14. Peixoto-Pino L, Barcala-Furelos R, Paz-García B, Varela-Casal C, Lorenzo-Martínez M, Gómez-Silva A, et al. The "DrownSafe" project: assessing the feasibility of a puppet show in teaching drowning prevention to children and parents. Children (Basel). 2024;11(1):19.

15. Kamstra P, Cook BR, Brander R, Lawes JC, Calverley H, Strugnell G, et al. Evaluating the impact of skill development for drowning prevention: a relationship-building approach to community engagement. Inj Prev. 2023;29(5):425-430.

16. Büyük F, Çamcı M. Retrospective evaluation of patients admitted to the emergency department due to drowning. Interdiscip Med J. 2024;15(52):48-53.

17. Szpilman D, Morgan PJ. Management for the drowning patient. Chest. 2021;159(4):1473-1483.

18. Bellini T, Brisca G, D'Alessandro M, Tibaldi J, Andreottola V, Conti C, et al. Association between early radiographic chest findings and clinical outcomes in paediatric drowning: a retrospective study in a tertiary Italian hospital. Eur J Pediatr. 2025;184(2):187.

19. Reizine F, Delbove A, Dos Santos A, Bodenes L, Bouju P, Fillâtre P, et al. Clinical spectrum and risk factors for mortality among seawater and freshwater critically ill drowning patients. Crit Care. 2021;25(1):372.

20. Şık N, Şenol HB, Çağlar A, Yılmaz D, Duman M. Early application of non-invasive ventilation for children with pulmonary edema after drowning. Pediatr Int. 2022;64(1):e14858.

21. Capela JB, Reis M, Almeida I, Gaspar L, Scortenschi E, Soares M, et al. Analysis of submersion accidents at a level II hospital and literature review. Paediatr Oncall J. 2024;21:93-97.

22. Dezfulian C, McCallin TE, Bierens J, Dunne CL, Idris AH, Kiragu A, et al. 2024 American Heart Association and American Academy of Pediatrics focused update on special circumstances: resuscitation following drowning. Circulation. 2024;150:e501-e516.

23. Jeswani NL, Khilji MF, Rizvi S, Al Reesi A. Epidemiology of drowning incidents among children at Sultan Qaboos University Hospital Oman. Oman Med J. 2021;36(6):e320.

24. Sacco MA, Aquila I. Post mortem molecular biomarkers of asphyxia: a literature review. Int J Mol Sci. 2024;25(21):11607.

25. Peri F, De Nardi L, Canuto A, Gaiero A, Noli S, Ferretti M, et al. Drowning in children and predictive parameters: a 15-year multicentre retrospective analysis. Pediatr Emerg Care. 2023;39(7):516-523.

26. Cousin VL, Pittet LF. Microbiological features of drowning-associated pneumonia: a systematic review and meta-analysis. Ann Intensive Care. 2024;14(1):61.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):376-386
DOI: 10.4274/BMB.galenos.2026.16046

Association of Tumor Distribution Pattern with Stage at Diagnosis and Presentation Pattern in Patients with Invasive Breast Cancer

İnvaziv Meme Kanserli Hastalarda Tümör Dağılım Paterni ile Tanı Anındaki Evre ve Başvuru Şekli Arasındaki İlişki

Erdoğan Şeyran, Azer Gökmen

University of Health Sciences Turkey, Van Training and Research Hospital, Department of Medical Oncology, Van, Turkey

Abstract

Objective: Breast cancer remains the most frequently diagnosed malignancy among women worldwide. Although the prognostic significance of tumor location has been extensively investigated, its association with stage at diagnosis and presentation pattern has been less well characterized. This study aimed to evaluate the association of tumor distribution pattern with stage at diagnosis and presentation pattern in women with invasive breast cancer.

Method: This retrospective single-center study included 238 women diagnosed with invasive breast cancer between January 2015 and January 2025. Clinicopathological data were obtained from institutional medical records. Associations between tumor distribution pattern and both stage at diagnosis and presentation pattern were evaluated using Pearson's chi-square test. Multivariable binary logistic regression analyses were performed to identify factors independently associated with advanced-stage disease (Stage III-IV) and symptomatic presentation.

Results: A total of 238 patients were included (mean age, 49.2±12.0 years), of whom 79.8% presented symptomatically. The upper outer quadrant was the most common tumor distribution pattern (47.5%). Tumor distribution pattern was significantly associated with both stage at diagnosis (χ^2=28.756, p=0.017; Cramér's V=0.20) and presentation pattern (χ^2=12.236, p=0.032; Cramér's V=0.23), although both associations were modest in magnitude. In multivariable analysis, tumors located in the upper inner quadrant [odds ratio (OR) 0.33, 95% confidence interval (CI) 0.14-0.82; p=0.017] and lower outer quadrant (OR 0.36, 95% CI 0.14-0.96; p=0.041) were

Öz

Amaç: Meme kanseri, dünya genelinde kadınlarda en sık tanı alan malignite olmaya devam etmektedir. Tümör yerleşiminin prognostik önemi kapsamlı olarak araştırılmış olmakla birlikte, tanı anındaki evre ve başvuru şekli ile ilişkisi daha az incelenmiştir. Bu çalışma, invaziv meme kanserli kadınlarda tümör dağılım paterni ile tanı anındaki evre ve başvuru şekli arasındaki ilişkiyi değerlendirmeyi amaçlamıştır.

Yöntem: Bu retrospektif, tek merkezli çalışmaya Ocak 2015 ile Ocak 2025 tarihleri arasında invaziv meme kanseri tanısı alan 238 kadın hasta dahil edildi. Klinikopatolojik veriler kurumsal tıbbi kayıtlardan elde edildi. Tümör dağılım paterni ile hem tanı anındaki evre hem de başvuru şekli arasındaki ilişkiler Pearson ki-kare testi kullanılarak değerlendirildi. İleri evre hastalık (Evre III-IV) ve semptomatik başvuru ile bağımsız olarak ilişkili faktörleri belirlemek amacıyla çok değişkenli ikili lojistik regresyon analizleri gerçekleştirildi.

Bulgular: Toplam 238 hasta çalışmaya dahil edildi (ortalama yaş, 49,2±12,0 yıl) ve hastaların %79,8'i semptomatik olarak başvurdu. Üst dış kadran en sık görülen tümör dağılım paterniydi (%47,5). Tümör dağılım paterni hem tanı anındaki evre (χ^2=28,756; p=0,017; Cramér's V=0,20) hem de başvuru şekli (χ^2=12,236; p=0,032; Cramér's V=0,23) ile anlamlı olarak ilişkiliydi; ancak her iki ilişkinin büyüklüğü de mütevazı düzeydeydi. Çok değişkenli analizde, üst iç kadran [olasılık oranı (OR)=0,33; %95 güven aralığı (GA): 0,14-0,82; p=0,017] ve alt dış kadran yerleşimli tümörler (OO=0,36; %95 GA: 0,14-0,96; p=0,041) ileri evre hastalık olasılığının daha düşük olmasıyla bağımsız olarak ilişkiliyken, HER2-zengin moleküler alt tip daha yüksek olasılıkla

Address for Correspondence: Erdoğan Şeyran, University of Health Sciences Turkey, Van Training and Research Hospital, Department of Medical Oncology, Van, Turkey
E-mail: erdogan.seyran@gmail.com **ORCID:** orcid.org/0000-0002-5238-0522
Received: 11.06.2026 **Accepted:** 21.08.2026 **Publication Date:** 29.09.2026

Cite this article as: Şeyran E, Gökmen A. Association of tumor distribution pattern with stage at diagnosis and presentation pattern in patients with invasive breast cancer. Bagcilar Med Bull. 2026;11(3):376-386

Abstract

independently associated with lower odds of advanced-stage disease, whereas the HER2-enriched molecular subtype was independently associated with higher odds (OR 8.16, 95% CI 1.75-38.02; p=0.008). For symptomatic presentation, T2-T4 tumors were the factor most strongly associated with symptomatic presentation (OR 5.98, 95% CI 2.48-14.40; p<0.001), whereas lower inner quadrant tumors remained independently associated with lower odds of symptomatic presentation (OR 0.16, 95% CI 0.04-0.62; p=0.008).

Conclusion: Tumor distribution pattern was independently associated with stage at diagnosis and presentation pattern; however, the observed associations were modest. After adjustment for established clinicopathological variables, specific tumor distribution patterns remained independently associated with advanced-stage disease and symptomatic presentation. Tumor burden, represented by T-stage, was the factor most strongly associated with symptomatic presentation. Overall, tumor distribution pattern may provide complementary clinicopathological information during the initial evaluation of invasive breast cancer but should be interpreted alongside established clinicopathological factors.

Keywords: Breast neoplasms, early detection of cancer, multicentric tumor, neoplasm staging, tumor localization

Öz

bağımsız olarak ilişkiliydi (OR=8,16; %95 GA: 1,75-38,02; p=0,008). Semptomatik başvuru açısından, T2-T4 tümörler semptomatik başvuru ile en güçlü ilişkili faktördü (OR=5,98; %95 GA: 2,48-14,40; p<0,001); alt iç kadran yerleşimli tümörler ise semptomatik başvuru olasılığının daha düşük olmasıyla bağımsız olarak ilişkili olmaya devam etti (OR=0,16; %95 GA: 0,04-0,62; p=0,008).

Sonuç: Tümör dağılım paterni tanı anındaki evre ve başvuru şekli ile bağımsız olarak ilişkiliydi; ancak gözlenen ilişkiler mütevazı düzeydeydi. Yerleşik klinikopatolojik değişkenlere göre düzeltme yapıldıktan sonra, belirli tümör dağılım paternleri ileri evre hastalık ve semptomatik başvuru ile bağımsız olarak ilişkili olmaya devam etti. Tümör yükünü yansıtan T-evresi, semptomatik başvuru ile en güçlü ilişkili faktördü. Genel olarak, tümör dağılım paterni invaziv meme kanserinin ilk değerlendirilmesi sırasında tamamlayıcı klinikopatolojik bilgi sağlayabilir; ancak yerleşik klinikopatolojik faktörlerle birlikte yorumlanmalıdır.

Anahtar kelimeler: Kanserin erken tanısı, meme neoplazmları, multisentrik tümör, neoplazi evrelemesi, tümör lokalizasyonu

Introduction

Breast cancer continues to pose a major global health challenge and remains the leading malignancy among women worldwide. Despite considerable progress in screening programs and therapeutic approaches, it still contributes substantially to cancer-related morbidity and mortality (1). In Turkey, breast cancer similarly represents the leading malignancy among women, and patients are frequently diagnosed at younger ages compared with many Western populations (2).

The biological behavior and clinical outcomes of breast cancer are influenced by numerous established clinicopathological factors, including tumor size, nodal involvement, histological subtype, tumor grade, molecular characteristics, and metastatic burden at diagnosis. Invasive ductal carcinoma constitutes the predominant histological subtype in routine clinical practice (3). Although many patients are identified at earlier stages, a considerable proportion still present with locally advanced or metastatic disease, particularly in regions where participation in screening programs remains limited or delays in medical evaluation are common (4).

Tumor distribution pattern within the breast, including both anatomical tumor location and multicentric disease, has increasingly attracted attention as a potential factor associated with tumor behavior and clinical outcomes. Among different anatomical regions, the upper outer quadrant is reported as the most frequent site of tumor development, which may be related to the relatively greater concentration of glandular breast tissue in this area (5). Beyond differences in incidence, several studies have suggested that tumor distribution pattern may also be associated with disease progression, stage at diagnosis, and prognosis. Particular interest has focused on tumors arising in the inner quadrants of the breast, as these lesions may demonstrate distinct lymphatic drainage characteristics involving the internal mammary chain (6-8). In addition, tumors involving the central breast region or multicentric disease have been associated with unique clinicopathological features and potentially different patterns of dissemination (9-11).

The mode of presentation at diagnosis also represents an important clinical issue in breast cancer management. In countries with widespread screening implementation, a greater proportion of tumors are detected before symptom onset. However, symptomatic presentation continues to predominate in many developing and middle-income regions and is frequently associated with more advanced disease at diagnosis (4,12). Delayed medical evaluation may therefore contribute to unfavorable stage distribution and poorer clinical outcomes (12,13).

Although the prognostic significance of breast tumor location has been extensively investigated, relatively few studies have evaluated whether tumor distribution pattern is independently associated with stage at diagnosis and presentation pattern after adjustment for important clinicopathological variables. Most previous studies have primarily focused on survival outcomes, whereas evidence regarding diagnostic pathways and stage distribution in real-world patient cohorts remains limited.

Accordingly, the present study was designed to investigate the association of tumor distribution pattern with stage at diagnosis and presentation pattern in patients with invasive breast cancer treated in a real-world clinical setting. In addition, multivariable logistic regression analyses were performed to determine whether these associations remained independent after adjustment for established clinicopathological factors, including tumor burden and intrinsic tumor biology. A better understanding of these relationships may contribute to improved interpretation of tumor distribution pattern as a clinically relevant characteristic during the diagnostic evaluation of invasive breast cancer.

Materials and Methods

Study Design and Participants

This retrospective observational study was conducted at a tertiary oncology center and included female patients diagnosed with invasive breast cancer between January 2015 and January 2025. The study was designed to evaluate the association of tumor distribution pattern with stage at diagnosis and presentation pattern. A total of 238 patients were eligible for inclusion. Women aged 18 years or older with histopathologically confirmed invasive breast cancer and complete clinicopathological records were included in the analysis. Patients with missing essential data or without histological confirmation of invasive disease were excluded. Sex was determined according to institutional medical records, and all participants were female.

Data Collection and Definitions

Clinical and pathological variables were extracted from institutional databases and archived medical records. Variables collected for analysis included age, menopausal status, tumor distribution pattern, histological subtype, histological grade, molecular subtype, TNM stage, and mode of presentation at diagnosis.

Tumor distribution pattern was categorized as upper outer quadrant, upper inner quadrant, lower outer quadrant, lower inner quadrant, central region, or multicentric disease. Multifocal tumors confined to a single breast quadrant were classified according to the corresponding anatomical quadrant, whereas tumors involving more than one breast quadrant were categorized as multicentric disease. Tumors located at the boundary between two adjacent quadrants were assigned to the quadrant containing the tumor epicenter, as determined from available imaging findings and pathological reports.

Disease stage was assigned according to the American Joint Committee on Cancer (AJCC) 8th edition TNM classification and categorized as stage I, II, III, or IV. For patients who received neoadjuvant therapy or presented with metastatic disease at diagnosis, stage was determined on the basis of clinical staging findings. For patients who underwent upfront surgery, pathological stage was used when available. Accordingly, throughout the manuscript, the term “stage at diagnosis” refers to the stage used for analysis based on this combined clinical-pathological staging approach.

Screening-detected cases were defined as asymptomatic tumors identified through routine mammographic screening, regardless of whether screening was performed within the national screening program or through individual opportunistic screening. Patients presenting with breast-related symptoms were classified as having symptomatic presentation.

Ethical Approval

Approval for the present study was granted by the Clinical Research Ethics Committee of University of Health Sciences Turkey, Van Training and Research Hospital (date: 16.01.2026; decision no: GOKAEK/2026-01-15). The study was carried out in compliance with the ethical principles of the Declaration of Helsinki. Given the retrospective design and use of existing medical records, informed consent was not required and was formally waived by the ethics committee.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 24.0 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean ± standard deviation or median (minimum-maximum), whereas categorical variables are summarized as frequencies and percentages.

Associations between tumor distribution pattern and stage at diagnosis, as well as presentation pattern, were initially evaluated using Pearson’s chi-square test. Monte Carlo

simulation was applied when appropriate to improve the reliability of probability estimation in contingency tables containing cells with low expected frequencies. The strength of statistically significant associations was quantified using Cramér's V coefficient. In addition, a sensitivity analysis excluding patients with multicentric disease was performed to evaluate whether the observed unadjusted associations between anatomical tumor distribution and both stage at diagnosis and presentation pattern persisted after removal of the multicentric subgroup.

To determine whether tumor distribution pattern was independently associated with advanced-stage disease and symptomatic presentation, two separate multivariable binary logistic regression models were constructed. The first model evaluated factors associated with advanced-stage disease (Stage III-IV vs. Stage I-II) and included age, menopausal status, tumor distribution pattern, histological grade, and molecular subtype. T and N stage were not included in this model because they are constituent components of the AJCC stage classification and their inclusion would have resulted in circular adjustment. The second model evaluated factors associated with symptomatic presentation (symptomatic vs. screening-detected) and included age, menopausal status, tumor distribution pattern, T-stage category (T2-T4 vs. T1), histological grade, histological subtype histological subtype [invasive ductal carcinoma (IDC) vs. non-IDC] and grouped molecular subtype. For the second model, molecular subtypes were grouped into hormone receptor-positive/HER2-negative (Luminal A and Luminal B/HER2-negative), HER2-positive (Luminal B/HER2-positive and HER2-enriched), and triple-negative categories to improve model stability and reduce sparse-data bias. Variables entered into both multivariable models were selected a priori based on their established clinical relevance rather than solely on univariable statistical significance.

Histological grade was unavailable for four patients; therefore, complete-case analysis was applied for multivariable regression. For the symptomatic-presentation model, patients with multicentric tumors were excluded because all multicentric tumors presented symptomatically, resulting in complete separation. In addition, two patients with Tx disease were excluded from this model because T-stage was included as an adjustment variable.

Adjusted odds ratios (ORs) with corresponding 95% confidence intervals (CIs) were calculated. Model fit was assessed using likelihood ratio chi-square statistics. All statistical tests were two-sided, and a p-value <0.05 was considered statistically significant. The methodology and reporting structure of the present study were developed in accordance with the STROBE Statement for observational studies.

Results

Patient Characteristics

A total of 238 women with histopathologically confirmed invasive breast cancer were included in the study. The mean age at diagnosis was 49.2±12.0 years (median, 46.5 years; range, 28-93 years). Premenopausal women constituted 49.2% of the cohort, while 43.7% were postmenopausal and 7.1% were perimenopausal. Most patients (79.8%) presented with breast-related symptoms, whereas 20.2% were diagnosed through screening.

The upper outer quadrant was the most common tumor distribution pattern (47.5%), followed by the upper inner quadrant (16.8%), lower outer quadrant (13.0%), lower inner quadrant (8.4%), central region (8.0%), and multicentric disease (6.3%). Invasive ductal carcinoma was the predominant histological subtype (67.6%). Histological grade was available for 234 patients, with grade 2 tumors representing the largest subgroup (50.9%). According to stage at diagnosis, 26.5% of patients had stage I disease, 33.6% stage II, 21.8% stage III, and 18.1% stage IV. Detailed demographic and clinicopathological characteristics are summarized in Table 1.

Association Between Tumor Distribution Pattern and Stage at Diagnosis

The distribution of stage at diagnosis differed significantly according to tumor distribution pattern (χ^2=28.756, p=0.017) (Table 2). However, the strength of this association was modest (Cramér's V=0.20). Patients with multicentric disease demonstrated the highest proportion of stage IV disease (53.3%), whereas tumors located in the upper inner and lower outer quadrants were more frequently diagnosed at earlier stages. The remaining tumor distribution patterns showed relatively balanced stage distributions across the four stage categories.

Association Between Tumor Distribution Pattern and Presentation Pattern

Tumor distribution pattern was also significantly associated with presentation pattern (χ^2=12.236, p=0.032) (Table 3). The strength of this association was modest (Cramér's V=0.23). All patients with multicentric disease

Table 1. Demographic and clinicopathological characteristics of the study population

Variable	Category	Value
Age at diagnosis (years)	Mean ± SD	49.2±12.0
	Median (range)	46.5 (28-93)
Menopausal status	Premenopausal	117 (49.2)
	Postmenopausal	104 (43.7)
	Perimenopausal	17 (7.1)
Presentation pattern	Screening-detected	48 (20.2)
	Symptomatic	190 (79.8)
Tumor distribution pattern	Upper outer quadrant	113 (47.5)
	Upper inner quadrant	40 (16.8)
	Lower outer quadrant	31 (13.0)
	Lower inner quadrant	20 (8.4)
	Central region	19 (8.0)
	Multicentric	15 (6.3)
Histological subtype	Invasive ductal carcinoma	161 (67.6)
	Invasive lobular carcinoma	20 (8.4)
	Mixed invasive carcinoma	14 (5.9)
	Other invasive carcinoma	27 (11.3)
	Other histological types	16 (6.7)
Histological grade*	Grade 1	20 (8.5)
	Grade 2	119 (50.9)
	Grade 3	95 (40.6)
Molecular subtype	Luminal A	50 (21.0)
	Luminal B/HER2-negative	90 (37.8)
	Luminal B/HER2-positive	44 (18.5)
	HER2-enriched	15 (6.3)
	Triple-negative	39 (16.4)
Tumor (T) stage	T1	97 (40.8)
	T2	105 (44.1)
	T3	20 (8.4)
	T4	14 (5.9)
	Tx	2 (0.8)
Nodal (N) stage	N0	104 (43.7)
	N1	64 (26.9)
	N2	42 (17.6)
	N3	28 (11.8)
Stage at diagnosis	Stage I	63 (26.5)
	Stage II	80 (33.6)
	Stage III	52 (21.8)
	Stage IV	43 (18.1)

Data are presented as mean ± standard deviation, median (range), or n (%). SD: Standard deviation, HER2: Human epidermal growth factor receptor 2. Other invasive carcinoma and other histological types represent distinct histopathological categories that were analyzed separately because of their different pathological characteristics.
*: Grade information was unavailable for four patients; percentages for histological grade were calculated based on the available data (n=234)

presented symptomatically. Among the anatomical tumor distribution patterns, lower inner quadrant tumors showed the highest proportion of screening-detected cases (40.0%), whereas symptomatic presentation remained the predominant mode of diagnosis across all other tumor distribution patterns. Sensitivity analyses excluding patients with multicentric disease demonstrated that the association between anatomical tumor distribution and stage at diagnosis was attenuated and no longer statistically significant (χ^2=14.56, df=12, p=0.266). Similarly, the association between anatomical tumor distribution and presentation pattern was also attenuated and did not remain statistically significant (χ^2=7.81, df=4, p=0.099). These findings indicate that the statistically significant unadjusted associations observed in the overall cohort were largely influenced by the multicentric subgroup.

Multivariable Analysis of Factors Associated with Advanced Stage at Diagnosis

To determine whether tumor distribution pattern remained independently associated with advanced-stage disease, a multivariable logistic regression model was constructed (Table 4). Compared with tumors located in the upper outer quadrant, tumors arising in the upper inner quadrant (OR, 0.33; 95% CI, 0.14-0.82; p=0.017) and lower outer quadrant (OR, 0.36; 95% CI, 0.14-0.96; p=0.041) were associated with significantly lower odds of presenting with stage III-IV disease.

Among molecular subtypes, HER2-enriched tumors demonstrated significantly higher odds of advanced-stage disease than Luminal A tumors (OR, 8.16; 95% CI, 1.75-38.02; p=0.008). Although multicentric disease was associated with increased odds of advanced-stage presentation (OR, 3.30; 95% CI, 0.93-11.73), this association did not reach statistical significance (p=0.066). Age, menopausal status, histological grade, and the remaining molecular subtypes were not independently associated with advanced-stage disease after adjustment. The overall regression model was statistically significant (Likelihood ratio χ^2=35.96, p=0.001). These findings indicate that selected anatomical tumor distribution patterns and tumor biology were independently associated with the odds of presenting with advanced-stage disease.

Multivariable Analysis of Factors Associated with Symptomatic Presentation

A second multivariable logistic regression model was performed to identify factors independently associated with symptomatic presentation (Table 5). Increasing age was

Table 2. Association between tumor distribution pattern and stage at diagnosis

Tumor distribution pattern	Stage I, n (%)	Stage II, n (%)	Stage III, n (%)	Stage IV, n (%)	Overall p-value
Upper outer quadrant	26 (23.0)	37 (32.7)	32 (28.3)	18 (15.9)	**0.017**
Upper inner quadrant	17 (42.5)	13 (32.5)	3 (7.5)	7 (17.5)	
Lower outer quadrant	7 (22.6)	15 (48.4)	6 (19.4)	3 (9.7)	
Lower inner quadrant	5 (25.0)	8 (40.0)	3 (15.0)	4 (20.0)	
Central region	6 (31.6)	5 (26.3)	5 (26.3)	3 (15.8)	
Multicentric disease	2 (13.3)	2 (13.3)	3 (20.0)	8 (53.3)	

Values are presented as n (row %). The association between tumor distribution pattern and stage at diagnosis was evaluated using Pearson's chi-square test (χ^2=28.756, degrees of freedom =15, p=0.017). The magnitude of the association was modest (Cramér's V=0.20)

Table 3. Association between tumor distribution pattern and mode of clinical presentation

Tumor distribution pattern	Screening-detected, n (%)	Symptomatic presentation, n (%)	Overall p-value
Upper outer quadrant	20 (17.7)	93 (82.3)	**0.032**
Upper inner quadrant	10 (25.0)	30 (75.0)	
Lower outer quadrant	4 (12.9)	27 (87.1)	
Lower inner quadrant	8 (40.0)	12 (60.0)	
Central region	6 (31.6)	13 (68.4)	
Multicentric disease	0 (0.0)	15 (100.0)	

Values are presented as n (row %). The association between tumor distribution pattern and mode of clinical presentation was evaluated using Pearson's chi-square test (χ^2=12.236, degrees of freedom =5, p=0.032). The magnitude of the association was modest (Cramér's V=0.23)

Table 4. Multivariable logistic regression analysis for factors associated with advanced stage at diagnosis

Variable	Adjusted OR	95% CI	p-value
Age at diagnosis, per year increase	0.98	0.94-1.02	0.301
Menopausal status			
Postmenopausal vs. Premenopausal	1.60	0.63-4.04	0.322
Perimenopausal vs. Premenopausal	1.02	0.32-3.22	0.979
Tumor distribution pattern			
Upper inner quadrant vs. Upper outer quadrant	**0.33**	**0.14-0.82**	**0.017**
Lower outer quadrant vs. Upper outer quadrant	**0.36**	**0.14-0.96**	**0.041**
Lower inner quadrant vs. Upper outer quadrant	0.64	0.21-1.98	0.441
Central region vs. Upper outer quadrant	0.82	0.27-2.48	0.730
Multicentric disease vs. Upper outer quadrant	3.30	0.93-11.73	0.066
Histological grade			
Grade 2 vs. Grade 1	1.62	0.46-5.68	0.454
Grade 3 vs. Grade 1	2.85	0.72-11.24	0.136
Molecular subtype			
Luminal B/HER2- vs. Luminal A	1.17	0.47-2.89	0.741
Luminal B/HER2+ vs. Luminal A	1.92	0.68-5.44	0.217
HER2-enriched vs. Luminal A	**8.16**	**1.75-38.02**	**0.008**
Triple-negative vs. Luminal A	0.63	0.19-2.08	0.445

Dependent variable: advanced stage at diagnosis (Stage III-IV vs. Stage I-II). Odds ratios (ORs) were estimated using multivariable logistic regression. Odds ratios were adjusted simultaneously for all variables included in the model. Reference categories were premenopausal status, upper outer quadrant location, Grade 1 histological grade and Luminal A subtype. Overall model fit: likelihood ratio χ^2=35.96, p=0.001. Four patients with missing histological grade were excluded; therefore, the analysis included 234 patients, CI: Confidence interval

Table 5. Multivariable logistic regression analysis for factors associated with symptomatic presentation

Variable	Adjusted OR	95% CI	p-value
Age at diagnosis, per year increase	**0.94**	**0.89-0.998**	**0.036**
Menopausal status			
Postmenopausal vs. Premenopausal	1.38	0.35-5.40	0.642
Perimenopausal vs. Premenopausal	**0.19**	**0.05-0.79**	**0.022**
Tumor distribution pattern			
Upper inner quadrant vs. Upper outer quadrant	0.75	0.25-2.24	0.609
Lower outer quadrant vs. Upper outer quadrant	1.28	0.33-4.98	0.719
Lower inner quadrant vs. Upper outer quadrant	**0.16**	**0.04-0.62**	**0.008**
Central region vs. Upper outer quadrant	0.46	0.11-1.88	0.278
T-stage category			
T2-T4 vs. T1	**5.98**	**2.48-14.40**	**<0.001**
Histological grade			
Grade 2 vs. Grade 1	1.84	0.52-6.53	0.345
Grade 3 vs. Grade 1	**4.68**	**1.04-21.00**	**0.044**
Histological subtype			
Non-IDC vs. IDC	0.70	0.29-1.68	0.424
Molecular subtype			
HER2-positive vs. HR-positive/HER2-negative	2.71	0.91-8.07	0.073
Triple-negative vs. HR-positive/HER2-negative	1.64	0.41-6.53	0.483

The dependent variable was symptomatic presentation versus screening-detected presentation. Odds ratios were estimated using multivariable binary logistic regression and were adjusted simultaneously for all variables included in the model. For multivariable analysis, molecular subtypes were grouped as hormone receptor-positive/HER2-negative (Luminal A and Luminal B/HER2-negative), HER2-positive (Luminal B/HER2-positive and HER2-enriched), and triple-negative. Histological subtype was categorized as invasive ductal carcinoma (IDC) versus non-IDC. Reference categories were premenopausal status, upper outer quadrant location, T1 stage, Grade 1, IDC, and HR-positive/HER2-negative molecular subtype. Multicentric tumors were excluded because all multicentric cases presented symptomatically, resulting in complete separation. Patients with missing histological grade or Tx disease were also excluded; therefore, the analysis included 217 patients. Overall model fit: likelihood ratio χ^2=67.05, p<0.001. OR: Odds ratio, CI: Confidence interval, HR: Hormone receptor, HER2: Human epidermal growth factor receptor 2

independently associated with lower odds of symptomatic presentation (OR, 0.94 per year; 95% CI, 0.89-1.00; p=0.036). Compared with premenopausal women, perimenopausal patients also had significantly lower odds of symptomatic presentation (OR, 0.19; 95% CI, 0.05-0.79; p=0.022), whereas postmenopausal status was not significantly associated with presentation pattern.

After adjustment for clinicopathological factors, patients with T2-T4 tumors had substantially higher odds of symptomatic presentation than those with T1 tumors (OR, 5.98; 95% CI, 2.48-14.40; p<0.001). Among tumor distribution patterns, lower inner quadrant tumors remained independently associated with lower odds of symptomatic presentation compared with upper outer quadrant tumors (OR, 0.16; 95% CI, 0.04-0.62; p=0.008). Grade 3 tumors were also independently associated with higher odds of symptomatic presentation than Grade 1 tumors (OR, 4.68; 95% CI, 1.04-21.00; p=0.044).

Neither histological subtype nor molecular subtype demonstrated statistically significant independent associations with symptomatic presentation after adjustment, although the HER2-positive group showed a trend toward higher odds of symptomatic presentation compared with the HR-positive/HER2-negative group (OR, 2.71; 95% CI, 0.91-8.07; p=0.073).

Patients with multicentric disease were excluded because all multicentric tumors presented symptomatically, resulting in complete separation. Patients with Tx disease were also excluded because T-stage was included in the multivariable model. The overall regression model was statistically significant (Likelihood ratio χ^2=67.05, p<0.001), indicating a statistically significant improvement over the intercept-only model.

Discussion

This retrospective study investigated the association of tumor distribution pattern with stage at diagnosis and presentation pattern in a real-world cohort of women with invasive breast cancer. The principal findings can be summarized as follows. First, tumor distribution pattern

was significantly associated with both stage at diagnosis and presentation pattern; however, the magnitude of these associations was modest. Second, after adjustment for relevant clinicopathological variables, tumors located in the upper inner and lower outer quadrants remained independently associated with lower odds of advanced-stage disease, whereas the HER2-enriched molecular subtype was independently associated with higher odds of presenting with stage III-IV disease. Third, T2-T4 tumors emerged as the strongest independent factor associated with symptomatic presentation. In addition, lower inner quadrant tumors remained independently associated with lower odds of symptomatic presentation after adjustment for T-stage and other clinicopathological variables, whereas Grade 3 tumors were associated with higher odds of symptomatic presentation. Collectively, these findings suggest that tumor distribution pattern may provide complementary clinicopathological information during the initial clinical assessment of patients with invasive breast cancer. Nevertheless, the observed effect sizes were modest, indicating that tumor distribution pattern should be interpreted alongside established clinicopathological factors rather than as an isolated determinant.

The association between breast tumor location and stage at diagnosis has been investigated in several previous studies; however, the reported findings have been inconsistent (14). Anatomically, the upper outer quadrant is the most common site of breast cancer development because of its relatively greater volume of glandular tissue, a finding that was also confirmed in our cohort (5). Beyond differences in tumor frequency, several studies have suggested that tumors arising in the inner quadrants may be associated with less favorable outcomes because of their lymphatic drainage to the internal mammary chain and the potential for occult nodal involvement (6-8,15). Other studies have reported distinct clinicopathological characteristics for centrally located and multicentric tumors, although their independent clinical significance remains uncertain (9-11,16,17). Consistent with these observations, our sensitivity analyses demonstrated that exclusion of multicentric tumors attenuated the associations between anatomical tumor distribution and both stage at diagnosis and presentation pattern. These findings suggest that the multicentric subgroup substantially influenced the statistically significant unadjusted associations observed in the overall cohort.

In the present study, tumor distribution pattern was significantly associated with stage at diagnosis; however, the magnitude of this association was modest. Furthermore, multivariable analysis demonstrated that tumors located in the upper inner and lower outer quadrants were independently associated with lower odds of advanced-stage disease compared with tumors in the upper outer quadrant, whereas multicentric disease showed only a non-significant trend toward higher odds of advanced-stage presentation after adjustment. These findings suggest that the relationship between tumor distribution pattern and stage at diagnosis is influenced by multiple clinicopathological factors rather than anatomical distribution alone (18). Although certain anatomical tumor distribution patterns remained independently associated with advanced-stage disease after adjustment, the modest effect sizes observed in the present study indicate that tumor distribution pattern should be interpreted as complementary to, rather than a substitute for, established clinicopathological factors.

An additional finding of the present study was the independent association between the HER2-enriched molecular subtype and advanced-stage disease at diagnosis. This finding is biologically plausible and is consistent with previous studies demonstrating the more aggressive clinical behavior of HER2-enriched tumors in the absence of HER2-targeted therapy, including higher proliferative activity and a greater likelihood of presenting with advanced disease. Although the introduction of contemporary anti-HER2 therapies has substantially improved clinical outcomes, our results suggest that intrinsic tumor biology remains an important factor associated with disease extent at diagnosis, independent of tumor distribution pattern. Therefore, assessment of molecular subtype together with anatomical tumor distribution may provide a more comprehensive characterization of disease at the time of diagnosis.

The relationship between tumor distribution pattern and presentation pattern has received considerably less attention than its association with survival outcomes. In general, breast cancers detected through organized screening programs are more likely to be diagnosed at earlier stages, whereas symptomatic presentation is commonly associated with greater tumor burden and more advanced disease (4,12,13). Consequently, factors associated with the timing or mode of detection may indirectly influence stage at diagnosis and subsequent treatment strategies.

In the present study, tumor distribution pattern was significantly associated with presentation pattern; however, the strength of this association was modest. After adjustment for relevant clinicopathological variables, T2-

T4 tumors emerged as the strongest factor independently associated with symptomatic presentation, emphasizing the predominant role of tumor burden in determining clinical detection. Nevertheless, lower inner quadrant tumors remained independently associated with lower odds of symptomatic presentation even after adjustment for T-stage, suggesting that this association cannot be fully explained by differences in tumor size at diagnosis alone. Grade 3 tumors were also independently associated with higher odds of symptomatic presentation, supporting the contribution of more aggressive tumor biology to clinically apparent disease.

The mechanisms underlying the observed association between lower inner quadrant tumors and presentation pattern remain uncertain. Given the modest effect size observed in the present study, this finding should be interpreted with caution and requires confirmation in larger prospective multicenter cohorts. Several explanations may account for this observation. Differences in healthcare utilization, participation in screening programs, breast anatomy, or other unmeasured clinicopathological characteristics may have contributed to the observed association. Because of the retrospective design and the possibility of residual confounding, the present findings should be considered hypothesis-generating rather than evidence of a causal relationship.

From a clinical perspective, the present findings suggest that tumor distribution pattern may provide complementary information during the initial evaluation of patients with invasive breast cancer (19). In the present study, tumor burden, represented by T-stage, emerged as the factor most strongly associated with symptomatic presentation, whereas tumor distribution pattern demonstrated additional independent associations beyond established clinicopathological variables. The persistence of the association between lower inner quadrant tumors and presentation pattern after adjustment for T-stage suggests that anatomical tumor distribution may be associated with clinical presentation independently of tumor size. However, given the modest effect sizes observed, tumor distribution pattern should be interpreted as a complementary clinicopathological characteristic rather than an independent determinant of clinical behavior.

Study Limitations

The present study has several strengths. It included a relatively large real-world cohort of patients with invasive breast cancer treated at a tertiary referral center over a 10-year period. Comprehensive clinicopathological data were available for most patients, allowing adjustment for important potential confounding factors using multivariable logistic regression analyses. In addition, the study evaluated both stage at diagnosis and presentation pattern, providing a broader assessment of the potential clinical relevance of tumor distribution pattern.

Nevertheless, several limitations should be acknowledged. First, the retrospective single-center design may have introduced selection bias and limits the generalizability of the findings. Second, residual confounding from unmeasured variables cannot be completely excluded despite multivariable adjustment. Detailed information regarding the regularity and frequency of screening mammography before diagnosis was not consistently available because of the retrospective design of the study. Therefore, the potential influence of prior screening adherence on presentation pattern could not be evaluated. Third, histological grade data were unavailable for a small number of patients, resulting in complete-case analysis for the multivariable regression models. In addition, stage at diagnosis was based on a combined clinical-pathological staging approach, which may have introduced some degree of classification heterogeneity between patients receiving neoadjuvant therapy and those undergoing upfront surgery.

The relatively limited number of screening-detected cases, together with the number of covariates included in the symptomatic-presentation model, may have reduced the precision of some adjusted estimates, as reflected by the wide CIs. Accordingly, some degree of model overfitting cannot be excluded. Furthermore, external validation of the present findings was not available, and survival outcomes were beyond the scope of this study; therefore, the prognostic implications of tumor distribution pattern could not be evaluated.

Finally, some tumor distribution subgroups, particularly multicentric, lower inner quadrant, and central tumors, contained relatively few patients. Consequently, subgroup-specific estimates should be considered exploratory and interpreted with caution. Moreover, the observed associations were modest in magnitude and therefore require confirmation in larger prospective multicenter studies before firm clinical conclusions can be drawn.

Conclusion

In conclusion, tumor distribution pattern was significantly associated with both stage at diagnosis and presentation

pattern in this real-world cohort of women with invasive breast cancer, although the observed associations were modest in magnitude. After adjustment for established clinicopathological variables, tumors located in the upper inner and lower outer quadrants remained independently associated with lower odds of advanced-stage disease, whereas lower inner quadrant tumors remained independently associated with lower odds of symptomatic presentation. Tumor burden, represented by T-stage, was the factor most strongly associated with symptomatic presentation. Overall, these findings suggest that tumor distribution pattern may provide complementary clinicopathological information during the initial evaluation of invasive breast cancer but should be interpreted together with established clinicopathological factors rather than as an independent determinant of clinical behavior. Further prospective multicenter studies incorporating survival outcomes are warranted to validate these findings and to further clarify the mechanisms underlying the observed associations.

Ethics

Ethics Committee Approval: Approval for the present study was granted by the Clinical Research Ethics Committee of University of Health Sciences Turkey, Van Training and Research Hospital (date: 16.01.2026; decision no: GOKAEK/2026-01-15).

Informed Consent: Given the retrospective design and use of existing medical records, informed consent was not required and was formally waived by the ethics committee.

Acknowledgments

Artificial intelligence-assisted language support (ChatGPT, OpenAI) was utilized during manuscript preparation. All scientific content was critically reviewed, verified, and approved by the authors, who assume full responsibility for the accuracy and integrity of the manuscript.

Footnotes

Authorship Contributions

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

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

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

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024;74(3):229-263.
2. Özmen V. Breast cancer in Turkey: clinical and histopathological characteristics (analysis of 13,240 patients). J Breast Health. 2014;10(2):98-105.
3. WHO Classification of Tumours Editorial Board. Breast tumours. 5th ed. Lyon: International Agency for Research on Cancer (IARC); 2019.
4. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin. 2025;75(1):10-45.
5. Bright CJ, Reulen RC, Winter DL, Stark DP, McCabe MG, Edgar AB, et al. Quadrant-specific breast cancer incidence and survival. Cancer Epidemiol. 2016;44:86-92.
6. Kong X, Xu F, Wang Z, Yang J, Chen J, Zhang X, et al. The prognostic impact of tumor location on breast cancer survival: a population-based analysis. JAMA Netw Open. 2020;3(10):e2016090.
7. Chang JH, Kim YJ, Lee JW, Kim HJ, Lee SB, Kim JS, et al. Tumor location and survival outcomes in breast cancer patients. Cancer Res Treat. 2017;49(3):689-697.
8. Noushi F, Spillane AJ, Uren RF, Gebski V, Akhurst T, Colletti PM, et al. Internal mammary node metastasis in breast cancer: implications for staging and treatment. Breast. 2011;20(2):109-115.
9. Hwang KT, Kim J, Jung J, Chang JH, Lee SB, Kim YA, et al. Impact of tumor location on breast cancer survival. Clin Breast Cancer. 2019;19(1):e27-e35.
10. Sohn VY, Arthurs ZM, Sebesta JA, Brown TA. Primary tumor location impacts breast cancer survival. Am J Surg. 2008;195(5):641-644.
11. Ji F, Xiao WK, Yang CQ, Wang YF, Tang ZH, Li J, et al. Association of tumor location with survival in breast cancer. Cancer Manag Res. 2019;11:2915-2925.
12. Pace LE, Mpunga T, Hategekimana V, Dusengimana JM, Habineza H, Bigirimana JB, et al. Delays in breast cancer presentation and diagnosis in low- and middle-income countries. Oncologist. 2015;20(7):780-788.
13. Malherbe F, Veller M. Palpable breast lumps: an age-based approach to evaluation and diagnosis. S Afr Fam Pract. 2022;64(1):e1-e7.
14. Shah A, Haider G, Abro N, Shaikh NA, Qureshi MA, Soomro S, et al. Correlation between tumor site and stage at presentation in breast cancer patients. Cureus. 2022;14(2):e22672.
15. Sarp S, Fioretta G, Verkooijen HM, Vlastos G, Rapiti E, Bouchardy C. Lower-inner quadrant breast cancers and prognosis. Ann Surg Oncol. 2007;14(3):1031-1039.
16. Poenaru MO, Amza M, Toma CV, Augustin FE, Pacu I, Zampieri G, et al. Multicentric and multifocal breast tumors—narrative literature review. Cancers (Basel). 2025;17(20):3380.

17. Avera E, Gansauer L, Yao K. Current understanding and distinct features of multifocal and multicentric breast cancers. Cancer Rep (Hoboken). 2023;6(12):e1851.

18. Gwak H, Jung SH, Suh YJ, Nam SJ, Han JH, Oh SJ, et al. Prognostic impact of multiple synchronous T1 breast cancer. Cancers (Basel). 2024;16(23):4019.

19. Fu WD, Wang XH, Lu KK, Lu YQ, Zhou JY, Huang QD, et al. Real-world outcomes for Chinese breast cancer patients with tumor location of central and nipple portion. Front Surg. 2022;9:993263.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):387-395
DOI: 10.4274/BMB.galenos.2026.97759

Factors Associated with Disability in Community-dwelling Oldest-old: Insights from Comprehensive Geriatric Assessment-based Nomogram

Seksen Yaş ve Üzeri Toplumda Yaşayan Yaşlılarda Disabilite ile İlişkili Faktörler: Kapsamlı Geriatrik Değerlendirmeye Dayalı Nomogram

Esra Çataltepe[1], Ayşe Fadıloğlu[2], Eda Çeker[3]

[1]University of Health Sciences Turkey, Konya Beyhekim Research and Training Hospital, Department of Geriatric Medicine, Konya, Turkey
[2]University of Health Sciences Turkey, Osmaniye Research and Training Hospital, Department of Geriatric Medicine, Osmaniye, Turkey
[3]University of Health Sciences Turkey, Ankara Etlik City Hospital, Department of Geriatric Medicine, Ankara, Turkey

Abstract

Objective: This study aimed to identify multidomain factors associated with disability in community-dwelling adults aged ≥80 years and to develop a nomogram estimating the probability of existing disability.

Method: This cross-sectional study included 385 community-dwelling adults aged ≥80 years. Disability was defined as dependency in at least one domain of activities of daily living (ADL) and/or instrumental activities of daily living (IADL). Candidate variables included cognition [mini-mental state examination (MMSE)], nutritional status (mini nutritional assessment-short form), depressive symptoms (geriatric depression scale), frailty (Fried phenotype), inflammatory status (C-reactive protein), number of medications, and demographic characteristics. Factors independently associated with disability were assessed with multivariable logistic regression in 371 participants with complete data. The nomogram was constructed from a five-variable model including age, MMSE, frailty, number of medications, and log(CRP). Discrimination was quantified by the area under the receiver operating characteristic curve (AUC), and internal validity was assessed through bootstrap resampling combined with calibration analysis.

Results: Disability prevalence was 74.8%. Disabled participants were older and had worse physical performance, lower cognitive scores, poorer nutritional status, higher depressive symptoms, and greater medication burden (all $p<0.05$). In multivariable analysis,

Öz

Amaç: Bu çalışmada, toplumda yaşayan 80 yaş ve üzerindeki bireylerde kapsamlı geriatrik değerlendirme bileşenlerini kullanarak disabilite ile ilişkili çok boyutlu faktörlerin belirlenmesi ve disabilite olasılığını tahmin etmeye yönelik bir nomogramın geliştirilmesi amaçlandı.

Yöntem: Kesitsel tasarımdaki çalışmaya toplumda yaşayan 80 yaş ve üzerindeki 385 birey dahil edildi. Disabilite, temel günlük yaşam aktiviteleri ve/veya enstrümantal günlük yaşam aktivitelerinde (EGYA) en az bir alanda bağımlılık olarak tanımlandı. Aday değişkenler arasında bilişsel durum (mini-mental durum muayenesi; MMSE), beslenme durumu (mini nütrisyonel değerlendirme-kısa form), depresif belirtiler (geriatrik depresyon ölçeği), kırılganlık (Fried kırılganlık fenotipi), C-reaktif protein, ilaç sayısı ve demografik özellikler yer aldı. Disabilite ile ilişkili bağımsız faktörler tam verisi bulunan 371 katılımcıda çok değişkenli lojistik regresyon analizi ile değerlendirildi. Nomogram; yaş, MMSE, kırılganlık, ilaç sayısı ve log(CRP) içeren beş değişkenli model temel alınarak oluşturuldu. Modelin ayırt edici gücü ROC eğrisi altında kalan alan (AUC) ile değerlendirildi; iç doğrulama bootstrap yeniden örnekleme yöntemi ve kalibrasyon analizi ile gerçekleştirildi.

Bulgular: Katılımcıların %74,8'inde disabilite saptandı. Disabilitesi olan bireyler daha ileri yaşta olup fiziksel performansları daha düşük, bilişsel fonksiyonları ve beslenme durumları daha kötü, depresif belirti düzeyleri ve kullandıkları ilaç sayısı daha yüksekti

Address for Correspondence: Esra Çataltepe, University of Health Sciences Turkey, Konya Beyhekim Research and Training Hospital, Department of Geriatric Medicine, Konya, Turkey
E-mail: esracataltepe@hotmail.com **ORCID:** orcid.org/0000-0003-1356-6700
Received: 07.07.2026 **Accepted:** 24.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Çataltepe E, Fadıloğlu A, Çeker E. Factors associated with disability in community-dwelling oldest-old: insights from comprehensive geriatric assessment-based nomogram. Bagcilar Med Bull. 2026;11(3):387-395

Abstract

lower MMSE scores [odds ratio (OR) 0.81, 95% confidence interval (CI) 0.73-0.89, p<0.001], frailty (OR 3.41, 95% CI 1.46-7.97, p=0.005), higher medication count (OR 1.25, 95% CI 1.10-1.41, p<0.001), and older age (OR 1.12, 95% CI 1.02-1.23, p=0.023) were independently associated with disability. A separate five-variable model including these four variables and log(CRP) was used to construct the nomogram and demonstrated good discrimination (AUC=0.818, 95% CI 0.775-0.856). Bootstrap validation indicated good calibration and minimal overfitting.

Conclusion: Cognitive impairment, frailty, medication burden, and older age were independently associated with disability in community-dwelling adults aged ≥80 years. The nomogram estimates the probability of existing disability rather than predicting future disability and requires external validation before clinical use.

Keywords: Activities of daily living, disability, frailty, geriatrics, older adults, nomogram

Öz

(tüm karşılaştırmalar için p<0,05). Çok değişkenli analizde düşük MMSE puanı [odds oranı (OR)=0,81; %95 güven aralığı (GA): 0,73-0,89; p<0,001], kırılganlık varlığı (OR=3,41; %95 GA: 1,46-7,97; p=0,005), daha fazla ilaç kullanımı (OR=1,25; %95 GA: 1,10-1,41; p<0,001) ve ileri yaş (OR=1,12; %95 GA: 1,02-1,23; p=0,023) disabilite ile bağımsız olarak ilişkili bulundu. Bu dört değişken ile log(CRP)'yi içeren ayrı bir beş değişkenli nomogram iyi düzeyde ayırt edici performans gösterdi (AUC=0,818; %95 GA: 0,775-0,856).

Sonuç: Toplumda yaşayan 80 yaş ve üzerindeki bireylerde bilişsel bozukluk, kırılganlık, ilaç yükü ve ileri yaş disabilite ile bağımsız olarak ilişkili bulundu. Nomogram, gelecekteki disabiliteyi öngörmek yerine mevcut disabilite olasılığını tahmin etmekte olup klinik kullanım öncesinde harici doğrulama gerektirmektedir.

Anahtar kelimeler: Disabilite, günlük yaşam aktiviteleri, kapsamlı geriatrik değerlendirme, kırılganlık, nomogram, yaşlı

Introduction

Population aging has led to a substantial increase in the number of adults aged ≥80 years, commonly referred to as the "oldest-old," who represent the fastest-growing segment of the global population (1). This age group is characterized by a high burden of chronic conditions (2), functional decline (3), and geriatric syndromes, making disability a major clinical and public health challenge (4).

Disability in older adults is increasingly recognized as a multifactorial condition arising from the interaction of multiple geriatric domains rather than a single disease process. Recent evidence highlights that cognitive impairment and physical frailty frequently coexist and synergistically increase the risk of functional decline and disability (5,6). Beyond these domains, treatment burden, often reflected by polypharmacy, has gained increasing recognition as a contributor to adverse outcomes in older adults (7). Similarly, systemic inflammation has been implicated as an underlying biological mechanism linking chronic disease burden, frailty, and functional decline (8).

Comprehensive geriatric assessment (CGA) provides an integrated framework to evaluate these interacting domains, enabling a holistic understanding of vulnerability in older adults. Disability, defined as dependence in activities of daily living (ADL) and instrumental activities of daily living (IADL), is strongly associated with increased morbidity and mortality in older adults (9,10). However, despite growing interest in multidimensional models, there remains limited evidence regarding the relative contributions of different geriatric domains within a single analytical framework, particularly among community-dwelling adults aged ≥80 years. Moreover, it remains unclear how domains such as cognition, frailty, inflammation, and treatment burden jointly relate to disability status in this age group

Accordingly, this study set out to identify the determinants of disability among community-dwelling adults 80 years of age and older through a CGA, with particular attention to how much each geriatric domain individually contributes to disability, and to develop a nomogram estimating the probability of existing disability

Materials and Methods

Study Design

This cross-sectional study was conducted at University of Health Sciences Turkey, Konya Beyhekim Research and Training Hospital. The study population consisted of community-dwelling adults aged ≥80 years enrolled in the Healthy Aging Program (YAŞAM). Within this program, participants undergo a CGA every 6 months. For the present study, only data obtained during each participant's first (index) home visit between January 2025 and January 2026 were included in the analysis. The CGA performed at this index visit included standardized assessments of cognitive function, nutritional status, depressive symptoms, physical performance, frailty, functional status, and medication use. Medical records of individuals aged ≥80 years who attended the YAŞAM outpatient clinic between January 2025 and January 2026 were retrospectively reviewed.

Clinical and functional data were obtained from routine CGAs performed during index visit. These assessments

included standardized tools evaluating cognitive function, nutritional status, depressive symptoms, physical performance, frailty, and medication use.

Laboratory data were obtained from blood samples collected on the same day as the index geriatric assessment and included routine biochemical parameters, including C-reactive protein (CRP). Due to its skewed distribution, CRP values were logarithmically transformed (log-CRP) before inclusion in statistical analyses.

Participants were excluded if they were unable to comply with the assessment procedures or had communication difficulties that precluded reliable evaluation during the CGA. Applying these criteria yielded a study sample of 385 participants. MMSE data were missing for five participants and CRP data for nine participants; these 14 participants were excluded from the multivariable and nomogram analyses, which were therefore performed as complete-case analyses in 371 participants (277 with and 94 without disability). Variable-specific missing data for descriptive characteristics are reported in Table 1.

CGA

All participants underwent a CGA as part of routine clinical care. Cognitive function was evaluated using the MMSE (11), nutritional status using the mini nutritional assessment-short form (MNA-SF) (12), and depressive symptoms with the geriatric depression scale (GDS) (13).

Functional status was assessed using ADL (14) and IADL (15). ADL encompassed six basic self-care activities: Bathing, dressing, toileting, transferring, continence, and feeding. IADL was assessed across all eight domains of the Lawton-Brody scale: Using the telephone, shopping, food preparation, housekeeping, laundry, transportation, responsibility for own medications, and handling finances. A participant was considered to have disability if dependence was present in at least one ADL or IADL domain.

Frailty was assessed using the Fried frailty phenotype (16), which includes unintentional weight loss, exhaustion, weakness, slow walking speed, and low physical activity. A participant was classified as frail if three or more of these criteria were met.

Gait performance was quantified with a 6-meter walk test, in which each participant walked this distance at a self-selected, comfortable pace while the elapsed time was captured with a stopwatch; dividing the distance by this time yielded walking speed in m/s.

A Takei handgrip dynamometer was used to quantify muscle strength, with participants seated and the elbow

Table 1. Baseline characteristics of participants according to disability status

Variable	Total (n=385)	Non-disabled (n=97)	Disabled (n=288)	p-value
Age, years (mean ± SD)	83.05±3.43	81.90±2.39	83.44±3.63	<0.001
Female sex, n (%)	238 (61.8)	49 (50.5)	189 (65.6)	0.008
BMI, kg/m² (mean ± SD)	28.39±5.43	27.30±4.76	28.77±5.60	0.021
Handgrip, kg (mean ± SD)	18.25±6.80	21.20±6.67	17.20±6.54	<0.001
Gait speed, m/s (mean ± SD)	0.80±0.31	0.98±0.29	0.74±0.29	<0.001
MMSE, median (IQR)	26 (22-29)	28 (27-30)	25 (19-28)	<0.001
MNA-SF, median (IQR)	12 (10-14)	13 (11-14)	12 (10-13)	<0.001
GDS score, median (IQR)	2 (1-5)	1 (0-3)	3 (1-6)	<0.001
Number of medications, median (IQR)	4 (3-6)	3 (2-5)	4 (3-6)	<0.001
Polypharmacy (≥5 medications), n (%)	158 (41.0)	28 (28.9)	130 (45.1)	0.005
Frailty (Fried phenotype; frail), n (%)	129 (33.5)	10 (10.3)	119 (41.3)	<0.001
Falls in the previous year, n (%)	128 (33.2)	27 (27.8)	101 (35.1)	0.19
Diabetes mellitus, n (%)	131 (34.0)	26 (26.8)	105 (36.5)	0.08
Hypertension, n (%)	286 (74.3)	70 (72.2)	216 (75.0)	0.58
Coronary artery disease, n (%)	113 (29.4)	22 (22.7)	91 (31.6)	0.10
CRP, mg/L, median (IQR)	3.78 (2.30-7.34)	3.08 (2.17-5.16)	4.10 (2.33-8.47)	0.018

BMI: Body mass index, MMSE: Mini-mental state examination, MNA-SF: Mini-nutritional assessment short form, GDS: Geriatric depression scale, CRP: C-reactive protein, SD: Standard deviation, IQR: Interquartile range. Data were available for BMI in 376 participants (97 non-disabled, 279 disabled), handgrip strength in 372 (97, 275), gait speed in 359 (95, 264), MMSE in 380 (97, 283), and CRP in 376 (94, 282); all other variables were complete. BMI and gait speed were compared with the independent-samples t-test, other continuous variables with the Mann-Whitney U test, and categorical variables with Pearson's chi-square test.

positioned at a 90-degree angle in line with standard protocol; the highest reading across repeated trials was retained for analysis.

Covariates

Patient demographic characteristics and comorbid conditions were obtained from the electronic medical records. Information on medication use was recorded, and polypharmacy was defined as the concurrent use of five or more medications (17).

Ethical Considerations

The study protocol was approved by the KTO Karatay University Ethics Committee (decision no: 2026/026, date: 30.04.2026), and institutional authorization was obtained from the host institution. The study was conducted in accordance with the Declaration of Helsinki. Because of the study's retrospective design, informed consent was waived.

Statistical Analyses

All statistical computations were carried out in SPSS (version 26.0; IBM Corp., Armonk, NY, USA), MedCalc (version 15.2; MedCalc Software Ltd., Ostend, Belgium), and R (version 4.4.2; R Foundation for Statistical Computing, Vienna, Austria). Depending on whether a continuous variable followed a normal distribution, its summary was reported either as the mean with standard deviation or as the median with the interquartile range; categorical variables were summarized as counts and proportions. Normality was checked with the Kolmogorov-Smirnov test. Group differences between participants with and without disability were examined with the independent-samples t-test when normality held, and with the Mann-Whitney U test otherwise; associations between categorical variables were tested with Pearson's chi-square test.

Factors independently associated with disability were identified through multivariable logistic regression (full model), in which candidate predictors were selected on the basis of clinical relevance together with statistical significance on univariate testing; effect estimates are presented as odds ratios (ORs) with corresponding 95% confidence intervals (CIs). Prior to model fitting, variance inflation factors were computed to rule out problematic collinearity among predictors; every value remained below 1.5, indicating that collinearity was not a concern for the multivariable model.

For the nomogram, a reduced model was fitted in the same complete-case sample (n=371), including the variables independently associated with disability in the full model (age, MMSE, frailty, and number of medications) together with log(CRP), which was retained because inflammation represented a clinically relevant biological domain in the multidimensional assessment of disability. The regression coefficients and intercept of this five-variable model (Table 3) were translated into a nomogram for individual-level estimation of the probability of existing disability. All performance measures reported below refer to this five-variable model. Model discrimination was quantified with the area under the receiver operating characteristic curve (AUC) with its 95% confidence interval, and the optimal probability cutoff was determined with the Youden index. The robustness of discrimination was checked internally through bootstrap resampling with 1000 iterations. Agreement between predicted and observed probabilities was examined graphically with a bootstrap-based calibration plot and with the Hosmer-Lemeshow goodness-of-fit test. As a sensitivity analysis, the models were re-estimated after excluding participants with CRP >10 mg/L, a level potentially indicative of acute-phase inflammation. A p-value below 0.05 (two-tailed) was taken as the threshold for statistical significance.

Results

The study sample comprised 385 community-dwelling adults aged 80 years or older, of whom 288 (74.8%) met criteria for disability. Compared with their non-disabled counterparts, disabled participants were, on average, older (83.44±3.63 vs. 81.90±2.39 years, $p<0.001$) and disproportionately female (65.6% vs. 50.5%, $p=0.008$). Mean BMI was slightly higher in the disabled group (28.77±5.60 vs. 27.30±4.76 kg/m², $p=0.021$). They had lower handgrip strength and gait speed, as well as poorer cognitive (MMSE), nutritional (MNA-SF), and depressive (GDS) scores (all $p<0.001$). Medication burden and polypharmacy were higher among disabled individuals ($p\leq0.005$), and frailty was more prevalent (41.3% vs. 10.3%, $p<0.001$). CRP levels were modestly higher in the disabled group ($p=0.018$), while no significant differences were observed for falls or comorbidities (Table 1).

In the multivariable logistic regression model (n=371), lower MMSE scores (OR 0.81, 95% CI 0.73-0.89, $p<0.001$), frailty (OR 3.41, 95% CI 1.46-7.97, $p=0.005$), higher medication count (OR 1.25, 95% CI 1.10-1.41, $p<0.001$), and older age (OR 1.12, 95% CI 1.02-1.23, $p=0.023$) were independently associated with disability. Nutritional status, depressive symptoms, sex, and log-CRP were not independently associated with disability (Table 2). In the sensitivity

analysis excluding participants with CRP >10 mg/L (n=72; remaining n=299), the association between log-CRP and disability was no longer observed (OR 0.99, 95% CI 0.58-1.69, p=0.97), while the associations for age, MMSE, frailty, and medication count remained materially unchanged.

Building on these results, a five-variable model including age, MMSE score, frailty status, medication count, and log(CRP) was fitted in the same sample (Table 3). In this model, all five variables were significantly associated with disability, including log(CRP) (OR 1.39, 95% CI 1.02-1.89, p=0.038); this association was no longer observed after excluding participants with CRP >10 mg/L (OR 1.08, 95% CI 0.64-1.81, p=0.78). The nomogram was constructed from this model (Figure 1). Each predictor contributed points in proportion to its regression coefficient, and summing these points across all five variables yielded each participant's estimated probability of existing disability, equivalent to 1/[1+exp(–LP)], where LP=-3.601+0.108 × age -0.226 × MMSE +1.188 × frailty (0/1) +0.227× number of medications +0.328 × ln(CRP).

Plotting the ROC curve of the five-variable model yielded an area under the curve (AUC) of 0.818 (95% CI, 0.775-0.856) (Figure 2), reflecting the model's ability to distinguish participants with disability from those without. Applying the Youden index identified 0.735 as the optimal probability cut-off, at which sensitivity was 69.3% (192/277) and specificity was 83.0% (78/94) (Youden's J =0.523; true positives 192, false negatives 85, true negatives 78, false positives 16).

Bootstrap resampling (1000 repetitions) yielded an optimism-corrected C-index of approximately 0.81, indicating minimal overfitting. Predicted and observed disability probabilities showed close agreement across the risk range (Figure 3), and the Hosmer-Lemeshow test indicated adequate calibration (χ^2=6.03, df=8, p=0.644).

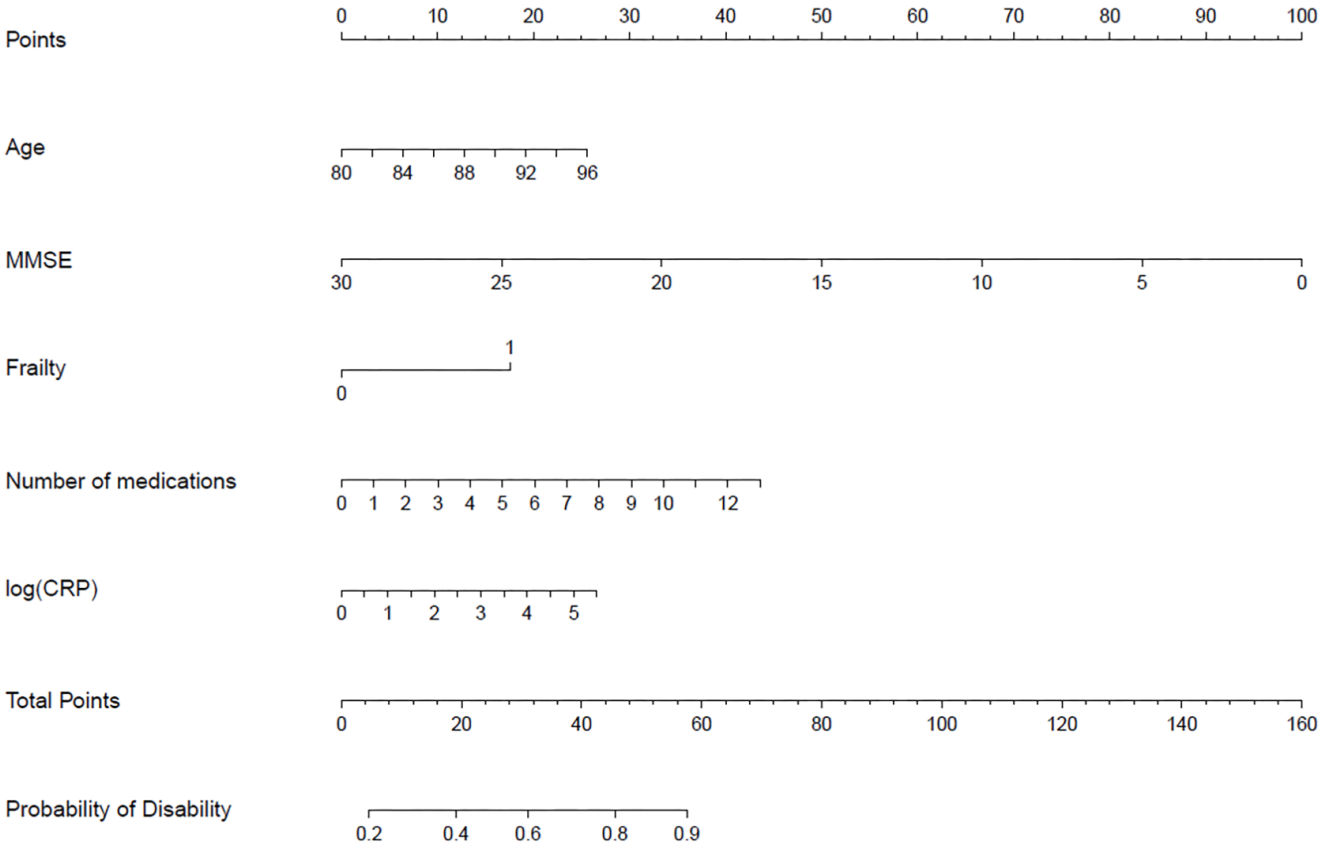

Figure 1. Nomogram based on the five-variable logistic regression model [age, MMSE, frailty, number of medications, and log(CRP); n=371] for estimating the probability of existing disability in community-dwelling adults aged 80 years and older. Frailty: 0= non-frail, 1= frail

MMSE: Mini-mental state examination, CRP: C-reactive protein

Table 2. Multivariable logistic regression for disability (full model; complete-case sample, n=371)

Variable	OR	95% CI	p-value
Age (per 1 year)	1.12	1.02-1.23	**0.023**
Sex (male)	0.84	0.48-1.47	0.55
MNA-SF	1.05	0.92-1.21	0.45
MMSE	0.81	0.73-0.89	**<0.001**
GDS	1.07	0.96-1.19	0.20
log(CRP)	1.35	0.99-1.85	0.056
Frailty (frail vs. non-frail)	3.41	1.46-7.97	**0.005**
Number of medications	1.25	1.10-1.41	**<0.001**

MNA-SF: Mini-nutritional assessment short form, MMSE: Mini-mental state examination, GDS: Geriatric depression scale, CRP: C-reactive protein, OR: Odds ratio, CI: Confidence interval

Table 3. Five-variable logistic regression model used to construct the nomogram (complete-case sample, n=371; 277 with and 94 without disability)

Variable	β (SE)	OR (95% CI)	p-value
Intercept	-3.601 (4.347)	—	0.407
Age (per 1 year)	0.108 (0.049)	1.11 (1.01-1.23)	**0.027**
MMSE (per 1 point)	-0.226 (0.048)	0.80 (0.73-0.88)	**<0.001**
Frailty (frail vs. non-frail)	1.188 (0.401)	3.28 (1.49-7.19)	**0.003**
Number of medications	0.227 (0.062)	1.25 (1.11-1.42)	**<0.001**
log(CRP)	0.328 (0.158)	1.39 (1.02-1.89)	**0.038**

MMSE: Mini-mental state examination, CRP: C-reactive protein (log = natural logarithm), SE: Standard error. Predicted probability of existing disability = 1/[1+exp(-LP)], where LP =-3.601+0.108 × age -0.226×MMSE +1.188× frailty (0/1) +0.227 × number of medications + 0.328 × ln(CRP). AUC 0.818 (95% CI 0.775–0.856); optimism-corrected C-index ≈0.81 (1000 bootstrap resamples); Hosmer-Lemeshow χ^2=6.03, df=8, p=0.644; Youden-optimal cut-off 0.735 (sensitivity 69.3%, specificity 83.0%)

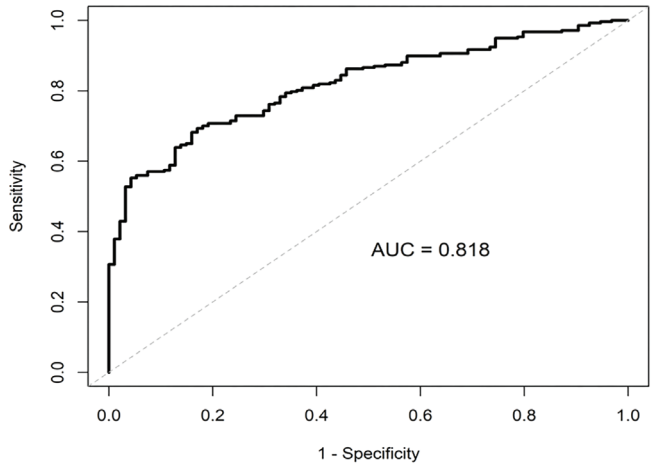

Figure 2. Receiver operating characteristic curve of the five-variable nomogram model for estimating the probability of existing disability (AUC 0.818, 95% confidence interval 0.775-0.856)

AUC: Area under the curve

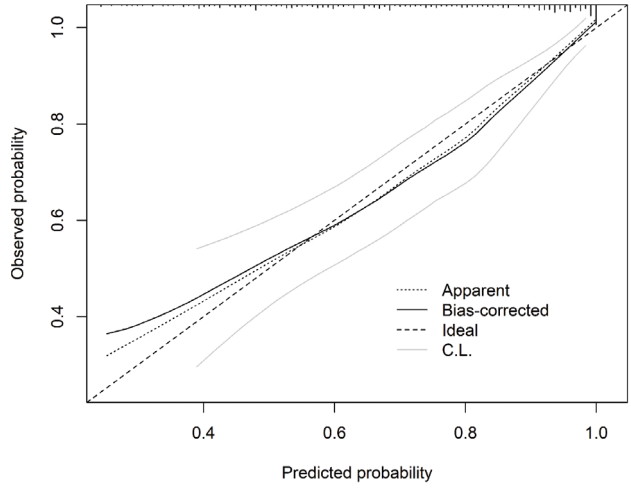

Figure 3. Bootstrap (1000 resamples) calibration plot of the five-variable nomogram model for estimating the probability of existing disability

C.L.: Confidence limits

Discussion

In this cohort of community-dwelling adults 80 years of age and older, disability was common, and independent associations emerged for cognitive performance, frailty status, medication burden, and age. Of these domains, cognition and frailty carried the greatest weight, while nutritional status and depressive symptoms were associated with disability in univariate analyses but lost statistical significance after full adjustment. These findings suggest that, in the oldest-old, disability is best understood as the cumulative expression of multidimensional vulnerability rather than as the consequence of a single isolated deficit.

The strong association between lower MMSE scores and disability in our cohort is clinically plausible and aligns with contemporary literature showing that cognitive impairment is closely linked to both basic and instrumental functional loss. Recent systematic evidence has shown that cognitive frailty and related cognitive deficits are associated with greater disability burden in community-dwelling older adults, particularly when cognitive decline coexists with reduced physical reserve (5,17). In practical terms, our findings support the notion that cognitive screening should not be viewed as ancillary for very old outpatients; rather, it should be considered a core component of disability risk assessment in the oldest-old.

Frailty was the other major independent determinant in our analysis. This is in line with a large body of evidence showing that frailty predicts incident and worsening ADL/IADL disability, and that the physical frailty phenotype captures frailty that is directly relevant to real-world function (18). Based on these findings, we developed and internally validated a practical nomogram incorporating age, MMSE, frailty status, medication burden, and log(CRP). The model demonstrated good discrimination (AUC =0.818) and good calibration after bootstrap validation, indicating reliable agreement between predicted and observed disability probabilities. This performance lies within the range reported for previously published disability prediction models in community-dwelling older adults (AUC 0.620–0.853)(19); however, most of those models predicted incident disability in longitudinal cohorts, whereas our model estimates the probability of existing disability, so this comparison should be interpreted with caution. Taken together, these data support frailty as a clinically meaningful and parsimonious indicator of disability in the oldest-old and suggest that the nomogram may help clinicians identify adults aged 80 years and older who are likely to have disability at the time of assessment.

Medication burden also remained independently associated with disability after multivariable adjustment. This finding is important because polypharmacy is often treated as a background characteristic rather than as an active determinant of poor outcomes. However, recent data suggest that polypharmacy is associated with incident disability in community-dwelling older adults and remains highly relevant in very old populations, even though the evidence base in those aged ≥85 years has historically been limited (20). Our results, therefore, reinforce the clinical relevance of medication review as part of the comprehensive assessment of disability. In a geriatric outpatient setting, the number of medications may function as a practical marker of treatment complexity, multimorbidity, and potential iatrogenic burden.

Another notable observation was that inflammatory burden, reflected by log-transformed CRP, was higher among disabled individuals in unadjusted analyses but was not independently associated with disability in the fully adjusted model. This pattern is also biologically plausible. Recent prospective evidence indicates that elevated high-sensitivity CRP is associated with a higher risk of ADL disability in older adults, but the magnitude of this association may be attenuated after adjustment for coexisting geriatric vulnerabilities such as frailty, sarcopenia-related performance deficits, and chronic disease burden (21). In our cohort, CRP likely captured part of the biological vulnerability underpinning disability, but it did not outweigh the explanatory value of cognition, frailty, and medication burden. Notably, the association was attenuated after excluding participants with markedly elevated CRP (>10 mg/L), suggesting that the relationship between CRP and disability was not robust across different CRP ranges. The same pattern was observed in the five-variable nomogram model, in which log(CRP) was statistically significant in the complete sample (p=0.038) but not after excluding these participants (p=0.78). This finding argues for cautious interpretation of CRP as an independent determinant of disability in this population.

By contrast, nutritional status and depressive symptoms did not remain independently associated with disability in the fully adjusted model. This does not imply that these domains are unimportant; rather, it suggests that their effects may be mediated through or overlap with stronger determinants such as frailty and cognition. In older adults, poor nutritional status commonly coexists with frailty and disability, while depressive symptoms may amplify functional decline indirectly through reduced motivation, lower activity, and poorer adherence. Yet, once core physical and cognitive domains are taken into account, their additional independent contribution may diminish. This interpretation is consistent with current multidomain geriatric frameworks, which emphasize the interconnected nature of these vulnerabilities.

Our study has several strengths. It focuses specifically on community-dwelling adults aged ≥80 years, a population that remains underrepresented in risk modeling studies. It also uses routinely collected CGA data, allowing simultaneous evaluation of several clinically relevant domains within the same analytical framework. Finally, we developed and internally validated a nomogram based on routinely collected comprehensive geriatric assessment variables, providing a practical tool for estimating the individual probability of existing disability.

Study Limitations

Several limitations should also be acknowledged. First, the cross-sectional design precludes causal inference; therefore, the identified determinants should be interpreted as correlates rather than predictors in a temporal sense, and the nomogram estimates the probability of existing disability rather than the risk of future disability. Second, disability was defined broadly as dependence in one or more ADL and/or IADL domains, which may have increased prevalence and reduced granularity. Third, our population consisted of attendees of a structured healthy

aging outpatient clinic, which may limit generalizability to institutionalized or more medically unstable older adults. Fourth, the multivariable analyses were performed as complete-case analyses after excluding 14 participants with missing MMSE or CRP data; although this represents a small proportion of the sample, some selection bias cannot be ruled out. Finally, although bootstrap internal validation demonstrated minimal optimism, external validation in independent oldest-old cohorts is required before routine clinical implementation.

Conclusion

In conclusion, our findings indicate that disability in community-dwelling oldest-old adults is independently associated with cognitive impairment, frailty, greater medication burden, and older age.

These results support the clinical value of CGA and suggest that disability status in the oldest-old is most meaningfully captured when cognitive, physical, and treatment-related domains are considered together rather than in isolation. The proposed nomogram may help estimate the probability of existing disability in this population; however, it was derived from cross-sectional data, should not be used to predict future disability, and requires external validation.

Ethics

Ethics Committee Approval: The study protocol was approved by the KTO Karatay University Ethics Committee (decision no: 2026/026, date: 30.04.2026), and institutional authorization was obtained from the host institution.

Informed Consent: In our study, data from patients between January 2025 and January 2026 were analyzed retrospectively.

Acknowledgments

For transparency, the authors note that an artificial intelligence-assisted language model (ChatGPT, OpenAI) was utilized to support text editing and language correction. This assistance was limited to linguistic refinement; all scientific content, critical analysis, and final editorial decisions were made exclusively by the authors.

Footnotes

Authorship Contributions

Concept: E.Ç., Design: E.Ç., A.F., Data Collection or Processing: E.Ç., Analysis or Interpretation: E.Ç., A.F., E.Çe., Literature Search: E.Ç., A.F., E.Çe., Writing: E.Ç., A.F., E.Çe.

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

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

References

1. Organization WH. Ageing and Health. [Internet]. 2025. Available from:https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.
2. Hajek A, König HH. Frequency and correlates of multimorbidity among the oldest old: study findings from the representative "survey on quality of life and subjective well-being of the very old in North Rhine-Westphalia (NRW80+)". Clin Interv Aging. 2023;18:41-48.
3. Dombrowsky T. Trajectories of functional decline in older adults: a latent class growth curve analysis. West J Nurs Res. 2023;45(8):715-725.
4. Berlau DJ, Corrada MM, Peltz CB, Kawas CH. Disability in the oldest-old: incidence and risk factors in the 90+ study. Am J Geriatr Psychiatry. 2012;20(2):159-168.
5. Tang KF, Teh PL, Lee SWH. Cognitive frailty and functional disability among community-dwelling older adults: a systematic review. Innov Aging. 2023;7(1):igad005.
6. Barrio-Cortes J, Benito-Sánchez B. Frailty and Disability in Older Adults. In: Bennett G, Goodall E, editors. The Palgrave Encyclopedia of Disability. Cham: Springer Nature Switzerland; 2024. p. 1-12.
7. Katsimpris A, Linseisen J, Meisinger C, Volaklis K. The association between polypharmacy and physical function in older adults: a systematic review. J Gen Intern Med. 2019;34(9):1865-1873.
8. Niu K, Hozawa A, Guo H, Kuriyama S, Ebihara S, Yang G, et al. Serum C-reactive protein even at very low (<1.0 mg/l) concentration is associated with physical performance in a community-based elderly population aged 70 years and over. Gerontology. 2008;54(5):260-267.
9. Ge Z, Li C, Li Y, Wang N, Hong Z. Lifestyle and ADL Are prioritized factors influencing all-cause mortality risk among oldest old: a population-based cohort study. Inquiry. 2024;61:469580241235755.
10. Gao Y, Du L, Cai J, Hu T. Effects of functional limitations and activities of daily living on the mortality of the older people: a cohort study in China. Front Public Health. 2023;10:1098794.
11. Güngen C, Ertan T, Eker E, Yaşar R, Engin F. Standardize mini mental test'in türk toplumunda hafif demans tanisinda geçerlik ve güvenilirliği [Reliability and validity of the standardized Mini Mental State Examination in the diagnosis of mild dementia in Turkish population]. Turk Psikiyatri Derg. 2002;13(4):273-281. Turkish.
12. Sarikaya D, Halil M, Kuyumcu ME, Kilic MK, Yesil Y, Kara O, 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.
13. Durmaz B, Soysal P, Ellidokuz H, Isik AT. Validity and reliability of geriatric depression scale-15 (short form) in Turkish older adults. North Clin Istanb. 2018;5(3):216-220.
14. Arik G, Varan HD, Yavuz BB, Karabulut E, Kara O, Kilic MK, et al. Validation of Katz index of independence in activities of

daily living in Turkish older adults. Arch Gerontol Geriatr. 2015;61(3):344-350.

15. 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.

16. Varan HD, Deniz O, Çöteli S, Doğrul RT, Kızılarslanoğlu MC, Göker B. Validity and reliability of Fried frailty phenotype in Turkish population. Turk J Med Sci. 2022;52(2):524-527.

17. Gnjidic D, Hilmer SN, Blyth FM, Naganathan V, Waite L, Seibel MJ, et al. Polypharmacy cutoff and outcomes: five or more medicines were used to identify community-dwelling older men at risk of different adverse outcomes. J Clin Epidemiol. 2012;65(9):989-995.

18. Kojima G. Frailty as a predictor of disabilities among community-dwelling older people: a systematic review and meta-analysis. Disabil Rehabil. 2017;39(19):1897-1908.

19. Zhang Z, Wang Q, Li Y, Zhou L. Disability risk prediction models in community-dwelling older adults: a systematic review. BMC Geriatrics. 2026;26(1):338.

20. Nishijima C, Harada K, Kurita S, Morikawa M, Fujii K, Kakita D, et al. Dietary variety and the relationship between polypharmacy and incident disability among Japanese community-dwelling older adults: a longitudinal study. Maturitas. 2025;193:108184.

21. Lai SM, Kuang L, Tang XL, Qiu CS, Huang HX, Liao DQ, et al. The role of high-sensitivity C-reactive protein in activities of daily living among middle-aged and older adults: a prospective cohort study. Nutrients. 2025;17(10):1732.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):396-403
DOI: 10.4274/BMB.galenos.2026.68553

Clinical Profile and Advanced Treatment Requirements in Children Admitted to the Pediatric Intensive Care Unit for Accidental Poisoning

Çocuk Yoğun Bakım Ünitesine Yatırılan Kazara Zehirlenme Olgularında Klinik Profil ve İleri Tedavi Gereksinimi

Kübra Boydağ Güvenç, Ebru Güney Şahin, İdris Abdullah Yılmaz, Fatih Varol, Cansu Durak

University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, Department of Pediatric Intensive Care, İstanbul, Turkey

Abstract

Objective: To describe the clinical profile, exposure patterns, treatments, and need for advanced pediatric intensive care unit (PICU)-level therapy among children admitted for accidental poisoning.

Method: This single-center, retrospective, descriptive study included children aged ≥1 month and <13 years who were admitted to the PICU for accidental poisoning between January 1, 2021, and January 31, 2026. Demographic characteristics, exposure agents and mechanisms, clinical findings, treatments, PICU length of stay, and hospital outcomes were obtained from electronic medical records. Severe clinical toxicity was defined as age-specific hypotension, a Glasgow Coma scale score of 8 or lower, clinically significant arrhythmia, seizure, or cardiac arrest. Invasive or non-invasive ventilation, vasopressor/inotropic support, continuous renal replacement therapy, therapeutic plasma exchange, extracorporeal membrane oxygenation, or hemoperfusion were considered advanced PICU-level therapies.

Results: Fifty-four patients were included. The median age was 29 months (interquartile range, 22-39.75), and 57.4% were male. Most exposures occurred at home (88.9%), resulted from exploratory ingestion (90.7%), and involved a single substance or product (98.1%). Medications were the most common exposure category (70.4%). Eighteen patients (33.3%) were symptomatic. Severe clinical toxicity occurred in three patients (5.6%), and five (9.3%) required advanced PICU-level therapy. Three patients received invasive

Öz

Amaç: Kazara zehirlenme nedeniyle çocuk yoğun bakım ünitesine (ÇYBÜ) yatırılan çocukların klinik özelliklerini, maruziyet örüntülerini, uygulanan tedavileri ve ÇYBÜ düzeyinde ileri tedavi gereksinimlerini değerlendirmek amaçlandı.

Yöntem: Bu tek merkezli, retrospektif ve tanımlayıcı çalışmaya, 1 Ocak 2021-31 Ocak 2026 tarihleri arasında kazara zehirlenme nedeniyle ÇYBÜ'ye yatırılan ≥1 ay ve <13 yaş arasındaki çocuklar dahil edildi. Demografik özellikler, maruziyet etkenleri ve mekanizmaları, klinik bulgular, tedaviler, ÇYBÜ yatış süresi ve hastane sonlanımları elektronik tıbbi kayıtlardan elde edildi. Ciddi klinik toksisite; yaşa göre hipotansiyon, Glasgow Koma skalası skorunun 8 veya altında olması, klinik olarak anlamlı aritmi, nöbet veya kardiyak arrest olarak tanımlandı. İnvaziv veya non-invaziv ventilasyon, vazopressör/inotrop desteği, sürekli renal replasman tedavisi, terapötik plazma değişimi, ekstrakorporeal membran oksijenasyonu veya hemoperfüzyon ileri ÇYBÜ tedavisi olarak kabul edildi.

Bulgular: Çalışmaya 54 hasta dahil edildi. Medyan yaş 29 ay (çeyrekler arası aralık, 22-39,75) ve hastaların %57,4'ü erkekti. Maruziyetlerin %88,9'u evde, %90,7'si keşif amaçlı kazara alım sonucunda gerçekleşti ve %98,1'i tek bir madde veya ürüne maruziyet şeklindeydi. İlaçlar en sık maruziyet grubuydu (%70,4). On sekiz hasta (%33,3) semptomatikti. Ciddi klinik toksisite üç hastada (%5,6), ileri ÇYBÜ tedavisi gereksinimi beş hastada (%9,3) saptandı. Üç hastaya invaziv mekanik ventilasyon, bir hastaya non-invaziv ventilasyon, üç hastaya vazopressör/inotrop desteği, iki hastaya sürekli renal

Address for Correspondence: Kübra Boydağ Güvenç, University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital, Department of Pediatric Intensive Care, İstanbul, Turkey
E-mail: kubrabydg@gmail.com **ORCID:** orcid.org/0000-0003-3881-6980
Received: 09.07.2026 **Accepted:** 03.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Boydağ Güvenç K, Güney Şahin E, Yılmaz İA, Varol F, Durak C. Clinical profile and advanced treatment requirements in children admitted to the pediatric intensive care unit for accidental poisoning. Bagcilar Med Bull. 2026;11(3):396-403

Abstract

mechanical ventilation, one received non-invasive ventilation, three received vasopressor/inotropic support, two underwent continuous renal replacement therapy, and two underwent therapeutic plasma exchange. The median PICU length of stay was 2 days, and all patients survived to hospital discharge.

Conclusion: Accidental poisonings predominantly occurred in young children, at home, and after single-agent exposure. Although advanced therapy was required infrequently, selected patients needed complex intensive care support. These findings emphasize the importance of safe storage practices and standardized triage approaches for determining the appropriate level of care.

Keywords: Accidental poisoning, advanced therapy, pediatric intensive care, pediatric toxicology, triage

Öz

replasman tedavisi ve iki hastaya terapötik plazma değişimi uygulandı. Medyan ÇYBÜ yatış süresi 2 gündü ve tüm hastalar sağ olarak taburcu edildi.

Sonuç: Kazara zehirlenmeler çoğunlukla küçük çocuklarda, ev ortamında ve tek ajan maruziyeti şeklinde gerçekleşmiştir. İleri tedavi gereksinimi düşük olmakla birlikte, seçilmiş hastalarda kompleks yoğun bakım desteği gerekmiştir. Bulgular, güvenli saklama önlemlerinin ve uygun bakım düzeyini belirleyecek standart triyaj yaklaşımlarının önemini göstermektedir.

Anahtar kelimeler: Çocuk yoğun bakım, ileri tedavi, kazara zehirlenme, pediyatrik toksikoloji, triyaj

Introduction

Accidental poisoning occurs most commonly in early childhood and typically takes place in the home. Young children are particularly vulnerable because of their natural tendency to explore their surroundings and place unfamiliar substances in their mouths. Medications and household chemicals are among the leading causes, although cleaning products, hydrocarbons, pesticides, rodenticides, and plant-derived products may also be involved (1,2). A substantial proportion of accidental exposures may be prevented through safe storage practices, appropriate medication dosing, and caregiver education (1,3).

Although most pediatric poisonings are asymptomatic or cause only mild clinical manifestations, some patients may develop altered mental status, respiratory depression, seizures, cardiac arrhythmias, or hemodynamic instability. These patients may require admission to the pediatric intensive care unit (PICU) for close monitoring, antidotal therapy, or respiratory and hemodynamic support (4). However, not every child admitted to the PICU requires an intensive care-level intervention. Some are admitted because of the characteristics of the ingested substance, the potential for delayed toxicity, or the need for close observation (4,5).

Previous studies have shown that many children admitted to the PICU after poisoning do not ultimately require PICU-level interventions (5,6). However, data comprehensively evaluating exposure characteristics, mechanisms of accidental poisoning, treatments administered, and the need for PICU-level support remain limited.

This study aimed to describe the demographic and clinical characteristics, exposure patterns, mechanisms of accidental poisoning, treatments administered, and short-term outcomes of children admitted to the PICU for accidental poisoning. As a secondary objective, we evaluated the proportion and clinical characteristics of patients who developed severe clinical toxicity or required advanced PICU-level therapy.

Materials and Methods

Study Design and Patient Population

This single-center, retrospective, descriptive study was conducted in the PICU of University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital. The electronic medical records of children aged 1 month to 12 years who were admitted to the PICU for accidental poisoning between January 1, 2021, and January 31, 2026, were reviewed retrospectively. Accidental poisoning was defined as unintentional exposure to medications, pesticides or rodenticides, chemical substances, or plant-derived products. Patients with intentional ingestions were excluded.

The study was approved by the Clinical Research Ethics Committee of University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital (decision no: 2026/114, date: 25.02.2026). The requirement for informed consent was waived by the ethics committee because of the retrospective study design. The study was conducted in accordance with the ethical principles of the Declaration of Helsinki.

During the study period, there was no standardized PICU admission protocol specifically for poisoning cases. Decisions regarding PICU admission were based on the patient's clinical condition, the characteristics of the exposure, the potential risk of delayed toxicity, the need for monitoring, recommendations from the National Poison

Consultation Center (114) when deemed necessary, and the judgment of the attending clinician.

Data Collection and Classification

Data extracted from the medical records included age, sex, body weight, history of regular medication use, Pediatric Risk of Mortality III (PRISM III) score (7), location of exposure, number of agents involved, availability of an estimated dose, mechanism of accidental exposure, clinical findings at presentation, treatments administered, PICU length of stay, and hospital outcome.

Exposures were classified into four main categories: medications, pesticides/rodenticides, chemical products, and plant-derived products. Medications were further categorized according to their pharmacological class as antipsychotics, analgesics/antipyretics, cardiovascular agents, colchicine, non-steroidal anti-inflammatory drugs, antidepressants, antiepileptic drugs, anticoagulants, benzodiazepines, antihistamines, and iron preparations. Chemical products were classified as alcohol-containing products, corrosive or cleaning agents, and hydrocarbons or organic solvents, whereas plant-derived products were categorized as toxic plants and volatile or essential oils.

For tabulation, multiple-agent exposure was defined as exposure to more than one separate substance or product. A cardiovascular combination preparation was categorized as a single cardiovascular agent exposure, although its active components were described in the Supplementary Table 1 when clinically relevant.

The mechanism of accidental exposure was classified into two categories. Unsupervised access to and ingestion of a substance present in the child's environment was defined as exploratory accidental ingestion, whereas administration of the wrong medication or an incorrect dose by a caregiver was classified as a caregiver-related medication or dosing error.

Patients were considered symptomatic at presentation if at least one exposure-related clinical finding was documented. The assessed findings included altered level of consciousness, somnolence, vomiting, age-specific hypotension, seizures, clinically significant arrhythmias, and other neurological symptoms. Because symptom data were obtained retrospectively from medical records, they were not documented in a standardized manner for all patients.

Severe Clinical Toxicity and Advanced PICU-level Therapy

For the purposes of this study, severe clinical toxicity was defined as the presence of at least one of the following: Age-specific hypotension, a Glasgow Coma scale score of 8 or lower, clinically significant arrhythmia, seizure, or cardiac arrest. Hypotension was assessed according to age-specific lower limits for systolic blood pressure.

The need for advanced PICU-level therapy was defined as the use of at least one of the following interventions: invasive mechanical ventilation, non-invasive ventilation, vasopressor or inotropic support, continuous renal replacement therapy (CRRT), therapeutic plasma exchange (TPE), extracorporeal membrane oxygenation, or hemoperfusion.

Because the mode of antidote administration—single dose, repeated dosing, or continuous infusion—could not be reliably distinguished in all cases, antidote use was not included in the definition of advanced PICU-level therapy. Intravenous lipid emulsion was recorded separately as a toxicology-specific rescue therapy and was not used as an independent criterion for advanced PICU-level therapy.

Treatments and Outcomes

Recorded treatments included activated charcoal, gastric lavage, antidotes, intravenous lipid emulsion, invasive and non-invasive respiratory support, vasopressor or inotropic

Table 1. Demographic and exposure characteristics of the study population

Age, months, median (IQR)	29 (22-39.75)
Male sex, n (%)	31 (57.4)
Body weight, kg, median (IQR)	15.0 (12.0-17.75)
PRISM III score, median (IQR)	0 (0-0)
History of regular medication use, n (%)	6 (11.1)
Location of exposure, n (%)	
Home	48 (88.9)
Outdoor setting	5 (9.3)
Another household	1 (1.9)
Number of agents involved, n (%)	
Single-agent exposure	53 (98.1)
Multiple-agent exposure	1 (1.9)
Estimated ingested dose known, n (%)	14 (25.9)
Mechanism of unintentional exposure, n (%)	
Exploratory ingestion by the child	49 (90.7)
Caregiver-related dosing error	5 (9.3)

Data are presented as n (%) or median (interquartile range), unless otherwise indicated. PRISM III: Pediatric Risk of Mortality III, IQR: Interquartile range

therapy, CRRT, and TPE. The antidotes administered were N-acetylcysteine, vitamin K, and ethanol.

The descriptive outcomes included the distribution of exposure agents and mechanisms, clinical characteristics at presentation, treatments administered, PICU length of stay, and hospital outcomes. The secondary outcomes were severe clinical toxicity and the need for advanced PICU-level therapy.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 26.0 (IBM Corp., Armonk, NY, USA). The distribution of continuous variables was assessed using the Shapiro-Wilk test and visual inspection. Continuous variables with non-normal distributions were presented as medians with interquartile ranges (IQRs), while categorical variables were summarized as numbers and percentages. Because of the small number of patients with severe clinical toxicity or requiring advanced PICU-level therapy, no comparative inferential analyses or multivariable modeling were performed for these groups. Patients who received advanced therapy were therefore described individually according to their clinical characteristics.

Results

A total of 54 children admitted to the PICU for accidental poisoning were included in the study. The median age was 29 months (IQR, 22-39.75), the median body weight was 15.0 kg (IQR, 12.0-17.75), and the median PRISM III score was 0 (IQR, 0-0). Thirty-one patients (57.4%) were male, and six (11.1%) had a history of regular medication use.

Most exposures occurred at home (n=48, 88.9%), while five (9.3%) occurred outdoors and one (1.9%) in another household. Fifty-three patients (98.1%) had a single-agent exposure, whereas one patient (1.9%) was exposed to two agents. The estimated dose was known in 14 cases (25.9%). The mechanism of exposure was classified as exploratory accidental ingestion in 49 patients (90.7%) and caregiver-related medication or dosing error in five (9.3%) (Table 1).

Medications were the most common exposure category, accounting for 38 cases (70.4%). Antipsychotics and analgesics/antipyretics were each involved in eight patients (14.8%), followed by cardiovascular agents in five (9.3%), colchicine in four (7.4%), and non-steroidal anti-inflammatory drugs and antidepressants in three patients each (5.6%). Antiepileptic drugs were involved in two patients (3.7%), while anticoagulants, benzodiazepines, antihistamines, iron preparations, and salicylate plus an angiotensin-converting enzyme inhibitor were each involved in one patient (1.9%).

Pesticide or rodenticide exposures occurred in nine patients (16.7%), chemical product exposures in five (9.3%), and plant-derived product exposures in two (3.7%). Within the pesticide/rodenticide group, eight patients had been exposed to rodenticides with an unknown active ingredient, and one to an agricultural pesticide. Among chemical exposures, alcohol-containing products were the most common subgroup, affecting three patients (5.6%) (Table 2).

At presentation, 18 patients (33.3%) were symptomatic. The most frequently documented findings were altered level of consciousness, defined as a Glasgow Coma scale score below 15 (n=14, 25.9%), somnolence (n=9, 16.7%), and vomiting (n=8, 14.8%). Age-specific hypotension was present in three patients (5.6%), while no seizures or clinically significant arrhythmias were observed. Three patients had agitation,

Table 2. Distribution of exposure categories and toxic agents

Medications, n (%)	38 (70.4)
Antipsychotics	8 (14.8)
Analgesics/antipyretics	8 (14.8)
Cardiovascular agents	5 (9.3)
Antigout agents (colchicine)	4 (7.4)
Non-steroidal anti-inflammatory drugs	3 (5.6)
Antidepressants	3 (5.6)
Antiepileptic drugs	2 (3.7)
Anticoagulants	1 (1.9)
Benzodiazepines	1 (1.9)
Antihistamines	1 (1.9)
Vitamin/mineral preparation (iron)	1 (1.9)
Salicylate plus an angiotensin-converting enzyme inhibitor*	1 (1.9)
Pesticides/rodenticides, n (%)	9 (16.7)
Rodenticide, active ingredient unknown	8 (14.8)
Agricultural pesticide, active ingredient unknown	1 (1.9)
Chemical products, n (%)	5 (9.3)
Alcohol-containing products	3 (5.6)
Corrosive/household cleaning product	1 (1.9)
Hydrocarbon/organic solvent	1 (1.9)
Plant-derived products, n (%)	2 (3.7)
Cardiotoxic plant (oleander)	1 (1.9)
Essential oil (eucalyptus oil)	1 (1.9)

*: This patient represented the only multiple-agent exposure and was exposed to both a salicylate and an angiotensin-converting enzyme inhibitor; the patient was counted once as a combined exposure subgroup

numbness, or other neurological symptoms recorded in free-text form. Clinical findings overlapped, and the categories of altered level of consciousness and somnolence included some of the same patients.

Activated charcoal was administered to 41 patients (75.9%), and gastric lavage to 28 (51.9%). Eleven patients (20.4%) received an antidote, while five (9.3%) required advanced PICU-level therapy. Forty patients (74.1%) received neither an antidote nor advanced PICU-level therapy, and two received both.

Severe clinical toxicity was identified in three patients (5.6%). All three had age-specific hypotension, and one patient with olanzapine exposure also had a Glasgow Coma scale score of 8. No seizures, clinically significant arrhythmias, or cardiac arrests occurred. Invasive mechanical ventilation was provided to three patients, non-invasive ventilation to one, and vasopressor or inotropic support to three. One patient received intravenous lipid emulsion, two underwent CRRT, and two underwent TPE (Table 3).

The median PICU length of stay was 2 days (IQR, 2-3; range, 1-16 days). All patients survived to hospital discharge.

The five patients who required advanced PICU-level therapy were 16-84 months of age (Supplementary Table 1). The patient exposed to methylated spirits received ethanol, invasive mechanical ventilation, vasopressor support, and CRRT; the indication for CRRT was severe metabolic acidosis and hemodynamic instability. The patient with olanzapine exposure required invasive mechanical ventilation because of depressed consciousness with a Glasgow Coma scale score of 8 and received vasopressor support for age-specific hypotension. The patient exposed to an agricultural pesticide with an unidentified active ingredient received invasive mechanical ventilation because of a decreased level of consciousness. In the patient exposed to a cardiovascular combination preparation containing a calcium channel blocker and an angiotensin-converting enzyme inhibitor, severe hemodynamic instability was managed with non-invasive ventilation, vasopressor support, intravenous lipid emulsion, CRRT, and TPE. The patient with paracetamol exposure underwent TPE because of hepatotoxicity and coagulopathy. PICU length of stay among these five patients ranged from 4 to 16 days.

Discussion

In this study, most children admitted to the PICU for accidental poisoning were of preschool age, and exposures occurred predominantly at home and involved a single agent. Medications were the most common exposure category. By contrast, severe clinical toxicity and the need for advanced PICU-level therapy were observed in only a small proportion of patients, and all survived to hospital discharge. These findings suggest that, although most accidental poisonings in children follow a

Table 3. Clinical presentation, management, and short-term outcomes

Variable	Value
Clinical presentation	
Symptomatic at presentation, n (%)	18 (33.3)
Altered level of consciousness (GCS <15), n (%)	14 (25.9)
Somnolence, n (%)	9 (16.7)
Vomiting, n (%)	8 (14.8)
Age-specific hypotension, n (%)	3 (5.6)
Seizure, n (%)	0 (0.0)
Clinically significant arrhythmia, n (%)	0 (0.0)
Other neurological symptoms, n (%)	3 (5.6)
Gastrointestinal decontamination and antidotal therapy	
Activated charcoal administration, n (%)	41 (75.9)
Gastric lavage, n (%)	28 (51.9)
Any antidote administered, n (%)	11 (20.4)
N-acetylcysteine, n	9
Vitamin K, n	1
Ethanol, n	1
Severe toxicity and advanced PICU-level therapy	
Severe clinical toxicity, n (%)	3 (5.6)
Advanced PICU-level therapy, n (%)	5 (9.3)
Invasive mechanical ventilation, n (%)	3 (5.6)
Non-invasive ventilation, n (%)	1 (1.9)
Vasopressor/inotropic support, n (%)	3 (5.6)
Intravenous lipid emulsion, n (%)	1 (1.9)
Continuous renal replacement therapy, n (%)	2 (3.7)
Therapeutic plasma exchange, n (%)	2 (3.7)
Short-term outcomes	
PICU length of stay, days, median (IQR)	2 (2-3)
PICU length of stay, range, days	1-16
Discharged alive, n (%)	54 (100.0)
In-hospital mortality, n (%)	0 (0.0)

Clinical findings were not mutually exclusive. Severe clinical toxicity was defined as age-specific hypotension, Glasgow Coma scale score ≤8, clinically significant arrhythmia, seizure or cardiac arrest. Advanced PICU-level therapy included invasive or non-invasive ventilation, vasopressor/inotropic support, continuous renal replacement therapy, therapeutic plasma exchange, extracorporeal membrane oxygenation, or hemoperfusion. Intravenous lipid emulsion was recorded separately as a toxicology-specific rescue therapy. GCS: Glasgow Coma scale, IQR: Interquartile range, PICU: Pediatric intensive care unit

relatively mild clinical course, selected patients may still require mechanical ventilation, vasopressor support, or extracorporeal therapies.

The median age of 29 months and the predominance of exploratory accidental ingestion in our cohort are consistent with previous studies showing that unintentional poisonings occur most frequently in early childhood (8,9).

In the series by Olguin et al. (8), which focused on unintentional medication poisonings, the median age was similarly 2 years, and most patients were younger than 5 years. This age-related predominance may reflect young children's tendency to explore their surroundings orally and their unsupervised access to potentially toxic substances (8). By restricting our cohort to children aged ≥1 month and <13 years with accidental exposures, we reduced the clinical and behavioral heterogeneity associated with intentional ingestions during adolescence.

The predominance of home-based exposures in our cohort (88.9%) is consistent with previous reports emphasizing the importance of access to toxic substances within the household in childhood accidental poisoning (2). Medications were the leading exposure category, accounting for 70.4% of cases, which is also in keeping with the medication-focused series reported by Olguin et al. (8). Differences in the distribution of medication classes across studies may reflect variations in study period, prescribing and medication-use patterns, and the restriction of our cohort to patients admitted to the PICU. However, given the small number of patients within individual exposure subgroups, no reliable association can be drawn between specific drug classes and a severe clinical course.

It is noteworthy that activated charcoal and gastric lavage were used frequently, even though only one-third of the patients were symptomatic at presentation. Current toxicology practice favors an individualized approach to gastrointestinal decontamination rather than its routine use, taking into account the ingested substance, estimated dose, formulation characteristics, time elapsed since exposure, and airway protection (10,11). Gastric lavage should therefore be reserved for highly selected situations in which the potential benefits are judged to outweigh the risks (11,12). The relatively high utilization rates observed in our cohort may reflect several factors, including local practice patterns, uncertainty regarding the amount ingested, concern about delayed or initially occult toxicity, and a precautionary approach in young children whose clinical assessment may be challenging. Some decontamination procedures may also have been initiated in the referring emergency department or another healthcare facility before admission to the PICU. However, the retrospective records did not consistently include the time elapsed since ingestion, the location at which the procedure was performed, the clinical rationale, or the estimated severity of the exposure. We were therefore unable to determine whether activated charcoal or gastric lavage was administered within an appropriate time window or according to a clearly documented toxicological indication. These findings should not be interpreted as demonstrating inappropriate treatment in individual patients; rather, they highlight the need for more consistent documentation and for locally standardized, indication-based gastrointestinal decontamination protocols aligned with contemporary toxicology recommendations.

The low proportion of patients requiring advanced PICU-level therapy (9.3%) suggests a potential discrepancy between PICU admission after acute poisoning and the subsequent need for intensive care interventions. Similarly, in the large cohort reported by Patel et al. (5), only approximately one in six children admitted to the PICU received a PICU-level intervention. Their age-specific models, based on clinical effects and selected exposure types, identified low-risk patients with a negative predictive value of approximately 95%. In the model's independent validation cohort, PICU admissions were projected to decrease by approximately one-third (5). In a subsequent prospective evaluation, the simplified clinical decision tool also demonstrated high sensitivity for identifying children at very low risk of clinically significant poisoning (6). The low rate of advanced PICU-level therapy observed in our smaller cohort, which was limited to accidental exposures, is broadly consistent with these findings. However, because time to presentation, detailed organ system findings, and local admission criteria could not be fully evaluated retrospectively, this similarity is insufficient to support the direct application of a triage model.

However, the low intervention rate should not be interpreted as evidence that these admissions were unnecessary. Defining intensive care solely in terms of invasive procedures or pharmacologic interventions may fail to capture other important components of care, including close monitoring and the early detection of clinical deterioration. Bateman (13) likewise noted that low intervention rates may reflect the scope of the intervention definition rather than inappropriate PICU admission. Accordingly, the main value of our findings lies not in

retrospectively classifying admissions as necessary or unnecessary, but in describing the clinical characteristics of patients who required advanced therapy and generating hypotheses for standardized triage approaches to be evaluated prospectively (5,6).

The diversity of causative agents among the five patients who required advanced therapy indicates that severe clinical deterioration was not confined to a single exposure category. The use of mechanical ventilation, vasopressor support, intravenous lipid emulsion, CRRT, and TPE exchange in these cases further illustrates that selected accidental poisonings may be associated with complex intensive care needs. Nevertheless, the fact that all patients survived to hospital discharge and that the median PICU length of stay was short suggests a favorable short-term prognosis in this cohort. However, the absence of mortality may reflect the small sample size, the restriction of the study to accidental exposures, and the limited number of high-risk toxic agents. These findings should therefore not be generalized to all pediatric poisoning cases.

Study Limitations

This study has several limitations. Its single-center, retrospective design introduces the possibility of inaccuracies related to the quality and completeness of medical records. Because the estimated dose was available for only approximately one-quarter of the patients, dose-response relationships could not be evaluated. Symptoms and clinical findings were not documented in a standardized manner across all cases, which may have resulted in misclassification or underreporting of symptom patterns. The time from exposure to presentation, the timing of decontamination, and the detailed rationale for PICU admission were not consistently recorded.

The absence of standardized, poisoning-specific PICU admission criteria represents an important source of selection bias. Admission decisions were influenced not only by the child's clinical condition, but also by the perceived toxicity of the agent, concern about delayed deterioration, the need for close monitoring, and individual clinician judgment. As a result, the study population may include both children admitted because of established clinical severity and clinically stable children admitted primarily as a precaution. Conversely, children with similar exposures who were managed in the emergency department or pediatric ward were not captured. Therefore, the observed rates of severe toxicity and advanced PICU-level therapy should be interpreted within the context of local admission practices and may not reflect the full spectrum of accidental poisoning in children.

Because the mode of antidote administration could not be reliably distinguished, antidote use was not included in the definition of advanced PICU-level therapy. The small number of patients with severe toxicity or requiring advanced therapy precluded comparative analyses.

Conclusion

Most children admitted to the PICU for accidental poisoning were young, had been exposed at home, and had single-agent exposures, with medications representing the most common causative category. Severe clinical toxicity and the need for advanced PICU-level therapy occurred in only a small proportion of patients, and all survived to hospital discharge. These findings reinforce the importance of safe storage practices in the home and highlight the need to review gastrointestinal decontamination practices and develop standardized triage approaches to identify the appropriate level of care for low-risk patients. However, given the retrospective design, no individual admission can be classified as unnecessary.

Ethics

Ethics Committee Approval: The study was approved by the Clinical Research Ethics Committee of University of Health Sciences Turkey, Sancaktepe Şehit Prof. Dr. İlhan Varank Training and Research Hospital (decision no: 2026/114, date: 25.02.2026).

Informed Consent: The requirement for informed consent was waived by the ethics committee because of the retrospective study design.

Acknowledgments

For transparency, the authors note that an artificial intelligence-assisted language model (ChatGPT, OpenAI) was utilized to support language correction. This assistance was limited to linguistic refinement; all scientific content, critical analysis, and final editorial decisions were made exclusively by the authors.

Footnotes

Authorship Contributions

Surgical and Medical Practices: K.B.G., E.G.Ş., İ.A.Y., F.V., C.D., Concept: K.B.G., F.V., C.D., Design: K.B.G., F.V., C.D., Data Collection or Processing: K.B.G., E.G.Ş., İ.A.Y., Analysis or Interpretation: K.B.G., F.V., C.D., Literature Search: K.B.G., E.G.Ş., İ.A.Y., Writing: K.B.G., E.G.Ş., İ.A.Y., F.V., C.D.

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

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

Supplementary Table Link: https://d2v96fxpocvxx.cloudfront.net/a426c3a3-a110-40af-a6dd-1b2b563ce9ac/content-images/346500ca-2933-4d7d-bb34-fbefdbfd9cc9.pdf

References

1. World Health Organization, UNICEF. World report on child injury prevention. 1st ed., Geneva: World Health Organization, 2008.
2. Meyer S, Eddleston M, Bailey B, Desel H, Gottschling S, Gortner L. Unintentional household poisoning in children. Klin Padiatr. 2007;219(5):254-270.
3. Lavon O, Ben-Zeev A, Bentur Y. Medication errors outside healthcare facilities: a national poison centre perspective. Basic Clin Pharmacol Toxicol. 2014;114(3):288-292.
4. Gupta D, Dhingra S. Toxicology in pediatric ICU. In: Brierley J, Tissières P, Deep A, editors. ESPNIC children's intensive care textbook. 1st ed. Cham: Springer; 2025. p. 801-819.
5. Patel MM, Travers CD, Stockwell JA, Numur EA, Geller RJ, Kamat PP, et al. Reducing childhood admissions to the PICU for poisoning (ReCAP2) by predicting unnecessary PICU admissions after acute intoxication. Pediatr Crit Care Med. 2018;19(2):e120-e129.
6. Grunwell JR, McCracken CE, Travers CD, Geller RJ, Kamat PP. Prospective evaluation of a clinical decision tool to reduce childhood admissions to PICUs for poisoning: ReCAP2. Clin Toxicol (Phila). 2019;57(12):1137-1141.
7. Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated pediatric risk of mortality score. Crit Care Med. 1996;24(5):743-752.
8. Olguin HJ, Garduño LB, Pérez JF, Pérez CF. Unintentional poisoning with drugs in a Mexican pediatric population. J Popul Ther Clin Pharmacol. 2011;18:e156-e160.
9. Gul H, Rashid N, Kamal M, Khan A, Ahmad M, Hussain I. Risk factors and outcomes of pediatric poisoning: a cross-sectional study from a tertiary care hospital in Peshawar, Pakistan. Cureus. 2025;17(5):e83463.
10. Hoegberg LCG, Gosselin S, Buckley NA, Wood DM, Shepherd G, Hanley J, et al. Recommendations from the clinical toxicology recommendations collaborative on the administration of activated charcoal in acute oral overdose. Clin Toxicol (Phila). 2026;64(6):419-475.
11. Gosselin S, Hoegberg LCG, Hoffman RS. Gut decontamination in the poisoned patient. Br J Clin Pharmacol. 2025;91(3):595-603.
12. Benson BE, Hoppu K, Troutman WG, Bedry R, Erdman A, Höjer J, et al.; American Academy of Clinical Toxicology; European Association of Poisons Centres and Clinical Toxicologists. Position paper update: gastric lavage for gastrointestinal decontamination. Clin Toxicol (Phila). 2013;51(3):140-146.
13. Bateman ST. Pediatric poisonings: do they really need that PICU bed? Pediatr Crit Care Med. 2017;18(7):727-728.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):404-413
DOI: 10.4274/BMB.galenos.2026.39206

Sarcopenia Risk Perception in Parkinson's Disease and Its Associations with Clinical Characteristics

Parkinson Hastalığında Sarkopeni Risk Algısı ve Klinik Özelliklerle İlişkisi

Hatice Betigül Meral[1], Abdülkadir Ermiş[2], Aygül Tantik Pak[2], Mehmet Yalman[2], Eren Toplutaş[3], Ece Akın[4], Sena Tolu[1]

[1]İstanbul Medipol University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, İstanbul, Turkey
[2]University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, Department of Neurology, İstanbul, Turkey
[3]İstanbul Medipol University Faculty of Medicine, Department of Neurology, İstanbul, Turkey
[4]İstanbul Medipol University Faculty of Medicine, İstanbul, Turkey

Abstract

Objective: This study aimed to evaluate sarcopenia risk perception in older adults with Parkinson's disease (PD) and to investigate its relationship with sarcopenia risk and clinical characteristics.

Method: This cross-sectional study included 52 older adults with PD and 52 age- and sex-matched healthy older adults. Demographic characteristics and clinical data, including disease duration, Hoehn and Yahr stage, and levodopa equivalent daily dose, were recorded. Sarcopenia risk was assessed using the mini sarcopenia risk assessment-7 (MSRA-7), whereas sarcopenia risk perception was evaluated using the Turkish version of the Sarcopenia Disease Risk Perception Scale (SDRPS-TR), which comprises the domains of perceived susceptibility and perceived severity. Correlations between variables were assessed using Spearman correlation analysis.

Results: Patients with PD had significantly lower MSRA-7 scores than healthy controls (30.19±7.19 vs. 33.65±5.98, p=0.028), and the prevalence of high sarcopenia risk was also significantly higher in this group (46.2% vs. 25.0%, p=0.041). However, no significant differences were observed between the groups regarding SDRPS-TR susceptibility, severity, or total scores (all p>0.05). Among patients with PD, age showed significant negative correlations with SDRPS-TR susceptibility, severity, and total scores, whereas educational level showed significant positive correlations with all SDRPS-TR domains (all p<0.05). In addition, living with family and participation in social activities were associated with higher SDRPS-TR susceptibility, severity, and total scores (all p<0.05).

Öz

Amaç: Bu çalışmanın amacı, Parkinson hastalığı (PH) olan yaşlı bireylerde sarkopeni risk algısını değerlendirmek ve bunun sarkopeni riski ve klinik özelliklerle ilişkisini araştırmaktır.

Yöntem: Bu kesitsel çalışmaya PH olan 52 yaşlı birey ile yaş ve cinsiyet açısından eşleştirilmiş 52 sağlıklı yaşlı birey dahil edildi. Hastalık süresi, Hoehn ve Yahr evresi ve levodopa eşdeğer günlük dozu dahil olmak üzere demografik özellikler ve klinik veriler kaydedildi. Sarkopeni riski mini sarkopeni risk değerlendirme ölçeği-7 (MSRA-7) ile, sarkopeni risk algısı ise algılanan duyarlılık ve algılanan ciddiyet alt boyutlarından oluşan sarkopeni hastalığı risk algısı ölçeği-Türkçe versiyonu (SDRPS-TR) kullanılarak değerlendirildi. Değişkenler arasındaki ilişkiler Spearman korelasyon analizi ile değerlendirildi.

Bulgular: PH'de MSRA-7 puanları sağlıklı kontrollere göre anlamlı olarak daha düşük bulundu (30,19±7,19'a karşı 33,65±5,98; p=0,028) ve bu grupta yüksek sarkopeni riski prevalansı da anlamlı olarak daha yüksekti (%46,2'ye karşı %25,0; p=0,041). Buna karşın SDRPS-TR duyarlılık, ciddiyet ve toplam puanları açısından gruplar arasında anlamlı farklılık saptanmadı (tüm p>0,05). PH grubunda, yaş ile SDRPS-TR duyarlılık, ciddiyet ve toplam puanları arasında anlamlı negatif; eğitim düzeyi ile tüm SDRPS-TR alanları arasında ise anlamlı pozitif ilişkiler bulundu (tümü için p<0,05). Ayrıca, ailesiyle yaşama ve sosyal aktivitelere katılım daha yüksek SDRPS-TR duyarlılık, ciddiyet ve toplam puanlarıyla ilişkiliydi (tümü için p<0,05).

Sonuç: PH olan yaşlı bireyler ile sağlıklı kontroller arasında sarkopeni risk algısı açısından anlamlı bir farklılık bulunmadı. İPH grubunda,

Address for Correspondence: Hatice Betigül Meral, İstanbul Medipol University Faculty of Medicine, Department of Physical Medicine and Rehabilitation, İstanbul, Turkey
E-mail: meralbetigul@gmail.com **ORCID:** orcid.org/0000-0002-7519-8184
Received: 15.07.2026 **Accepted:** 07.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Meral HB, Ermiş A, Tantik Pak A, Yalman M, Toplutaş E, Akın E, et al. Sarcopenia risk perception in Parkinson's disease and its associations with clinical characteristics. Bagcilar Med Bull. 2026;11(3):404-413

Abstract

Conclusion: Sarcopenia risk perception did not differ between older adults with PD and healthy controls. Among patients with PD, lower risk perception was associated with advanced age, whereas higher educational level, living with family, and participation in social activities were associated with higher risk perception. These findings highlight the importance of integrating patient education and awareness strategies into routine sarcopenia screening and management in patients with PD.

Keywords: Older adults, Parkinson's disease, risk perception, sarcopenia

Öz

ileri yaş daha düşük sarkopeni hastalığı risk algısıyla ilişkiliyken, daha yüksek eğitim düzeyi, ailesiyle yaşama ve sosyal aktivitelere katılım daha yüksek sarkopeni hastalığı risk algısıyla ilişkili bulundu. Bu bulgular, PH'de rutin sarkopeni tarama ve yönetim süreçlerine hasta eğitimi ve farkındalık stratejilerinin entegre edilmesinin önemini vurgulamaktadır.

Anahtar kelimeler: Parkinson hastalığı, risk algısı, sarkopeni, yaşlı bireyler

Introduction

Parkinson's disease (PD) is one of the most common neurodegenerative disorders worldwide after Alzheimer's disease and significantly impairs physical functioning and quality of life due to its progressive nature (1,2). The disease is characterized not only by motor manifestations such as bradykinesia, rigidity, and tremor, but also by a wide spectrum of non-motor symptoms including cognitive, autonomic, and sensory dysfunctions (3). In recent years, increasing evidence has demonstrated that PD is not limited to neurological involvement and may also exert substantial effects on the musculoskeletal system (4,5). Current evidence indicates that sarcopenia is more prevalent among individuals with PD compared with healthy older adults and may influence disease progression (6,7).

Sarcopenia is a progressive muscle disorder associated with aging and characterized by deterioration in muscle structure and physical performance (8). It has been recognized as an independent disease entity within the ICD-10-CM classification system and is considered a major geriatric syndrome. The updated diagnostic criteria proposed by the Asian Working Group for Sarcopenia recommend the combined assessment of muscle mass, muscle strength, and physical performance for the diagnosis of sarcopenia (9). Sarcopenia is associated with adverse clinical outcomes including increased fall risk, loss of functional independence, hospitalization, and mortality, particularly in older populations (10).

Recent studies in patients with PD have investigated body composition changes, physical performance, fall risk, non-motor symptoms, balance impairment, quality of life, and other clinical factors associated with sarcopenia (11-13). However, to the best of our knowledge, no previous study has specifically evaluated sarcopenia risk perception in patients with PD. Risk perception is an important cognitive process that influences health behaviors and engagement in preventive approaches (14). Therefore, assessing sarcopenia risk perception in patients with PD may contribute to early recognition and the development of preventive interventions.

The Turkish version of the Sarcopenia Disease Risk Perception Scale (SDRPS-TR), a disease-specific instrument designed to assess perceived susceptibility to and perceived severity of sarcopenia, was shown to be valid and reliable in the Turkish older adult population (15). The scale demonstrated high internal consistency and strong psychometric properties, supporting its use as a reliable tool for assessing sarcopenia awareness among older adults (15). Accordingly, the present study aimed to evaluate sarcopenia risk perception among older adults with PD using the SDRPS-TR and to investigate its association with sarcopenia risk status as well as clinical and demographic variables.

Materials and Methods

Study Participants

This cross-sectional study included patients with PD who were consecutively recruited from the outpatient clinics of the departments of physical medicine and rehabilitation and neurology. Patients with idiopathic PD fulfilled the Movement Disorder Society clinical diagnostic criteria (16). Age- and sex-matched healthy control participants were recruited from individuals undergoing routine health assessments at the same institution during the study period. The study protocol was approved by the Local Institutional Non-Interventional Clinical Research Ethics Committee of İstanbul Medipol University (decision no: 866, date: 14.05.2026) and all participants provided written informed consent before enrollment.

Eligible participants were ambulatory adults aged 65 years or older. Patients with PD were required to have mild-to-moderate disease severity (Hoehn and Yahr stage 1-3). General cognitive function was assessed in all participants using the mini-mental state examination (MMSE), and individuals with MMSE scores <24 were excluded (17).

For patients with PD, neurological disorders other than PD were excluded. For healthy controls, any neurological disorder was considered an exclusion criterion. A history of psychiatric disorders and severe medical conditions that could affect study participation or the interpretation of outcomes were exclusion criteria for both groups. These conditions included New York Heart Association class III-IV heart failure, dialysis-dependent end-stage renal disease, severe or very severe chronic obstructive pulmonary disease, severe osteoarthritis, and active malignancy. In addition, patients with secondary parkinsonism (e.g., drug-induced or trauma-related parkinsonism) were excluded.

Clinical Assessments

Demographic characteristics including age, gender, educational level, living status, and participation in social activities were recorded for all participants. Educational level was categorized into five groups: Illiterate (1), primary school (2), middle school (3), high school (4), and university (5). Clinical characteristics of patients with PD included disease duration and Hoehn and Yahr stage. Antiparkinsonian medication doses were recorded, and the Levodopa equivalent daily dose (LEDD) was calculated using the standardized conversion factors proposed by Tomlinson et al. (18). Comorbidities, including hypertension, diabetes mellitus, dyslipidemia, and coronary artery disease, were documented. Smoking status, history of alcohol consumption, history of falls during the previous year, and the number of medications used (polypharmacy defined as the concurrent use of >5 medications) were also recorded.

Sarcopenia Risk Assessment

Sarcopenia risk was assessed using the Turkish version of the mini sarcopenia risk assessment questionnaire-7 (MSRA-7). The MSRA questionnaire was originally developed by Rossi et al. (19) to screen for sarcopenia risk. The MSRA-7 consists of seven items evaluating age, recent hospitalization, physical activity level, dietary habits, weight loss, and mobility status. Total scores range from 0 to 40, with lower scores indicating a higher risk of sarcopenia. The Turkish version of the MSRA-7 has been shown to be a valid and reliable instrument for screening sarcopenia risk in older adults (20). In accordance with the original scoring system, a total score of ≤30 was considered indicative of increased sarcopenia risk.

Sarcopenia Disease Risk Perception Scale (SDRPS)

Sarcopenia risk perception was assessed using the SDRPS-TR. The original SDRPS was developed and validated by Zhang et al. (21) to evaluate individuals' perceptions regarding the risk of sarcopenia among older adults. The scale was subsequently translated, culturally adapted, and validated for use in the Turkish population, demonstrating excellent validity and reliability (15). The SDRPS-TR consists of 15 items organized into two subscales: perceived susceptibility (8 items) and perceived severity (7 items). Each item is scored on a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). Higher scores indicate a greater perception of sarcopenia risk.

Ethics Statement

The study protocol was approved by the Local Institutional Non-Interventional Clinical Research Ethics Committee of İstanbul Medipol University (decision no: 866, date: 14.05.2026). The study was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to enrollment.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics (version 26.0; IBM Corp., Armonk, NY, USA). Continuous variables were presented as mean ± standard deviation, and categorical variables as frequencies and percentages. The distribution of continuous variables was assessed using the Shapiro-Wilk test and visual inspection of histograms and Q-Q plots. For normally distributed variables, comparisons between groups were performed using the Student's t-test, whereas the Mann-Whitney U test was used for variables that were not normally distributed. Categorical variables were analyzed using the Pearson chi-square test or Fisher's exact test, as appropriate. The relationships between continuous variables that did not follow a normal distribution were analyzed using Spearman's rank correlation coefficient. All statistical tests were two-sided, and a p-value <0.05 was considered statistically significant.

Results

A total of 104 participants were included in the study, comprising 52 patients with PD and 52 age- and sex-

matched healthy controls. The demographic and clinical characteristics of the study population are presented in Table 1.

There were no significant differences between the PD and control groups regarding age, sex distribution, body mass index (BMI), education level, hypertension, type 2 diabetes, dyslipidemia, coronary artery disease, polypharmacy, current smoking, alcohol use, living status, or social activities (all p>0.05). However, a significantly higher proportion of patients with PD had experienced a fall during the previous year compared with controls (p=0.009). Among patients with PD, the mean disease duration was 7.97±5.61 years, the mean LEDD was 631.35±356.19 mg/day, and Hoehn and Yahr stage 2.5 was the most common disease stage.

As shown in Table 2, patients with PD had significantly lower MSRA-7 scores than healthy controls (30.19±7.19 vs. 33.65±5.98, p=0.028). In addition, the prevalence of high sarcopenia risk (MSRA-7 ≤30) was significantly higher in the PD group than in the control group (46.2% vs. 25.0%,

Table 1. Demographic and clinical characteristics of the study population

	Parkinson's disease (n=52)	Controls (n=52)	p-value
Age (years)	74.00±6.80	72.31±6.16	0.195
Gender n (%)			
Female	25 (48.1)	26 (50.0)	1.000
Male	27 (51.9)	26 (50.0)	
BMI (kg/m²)	26.21±4.18	25.52±3.70	0.513
Education level, n (%)			
Illiterate	10 (19.2)	9 (17.3)	0.506
Primary school	13 (25.0)	8 (15.4)	
Middle school	5 (9.6)	9 (17.3)	
High school	8 (15.4)	12 (23.1)	
University	16 (30.8)	14 (26.9)	
Hypertension, n (%)	21 (40.4)	29 (55.8)	0.169
Type 2 diabetes, n (%)	7 (13.5)	12 (23.1)	0.310
Dyslipidemia, n (%)	15 (28.8)	13 (25.0)	0.825
Coronary artery disease, n (%)	2 (3.8)	3 (5.8)	1.000
Polypharmacy (>5 drugs), n (%)	21 (40.4)	19 (36.5)	0.840
Current smoking, n (%)	3 (5.8)	9 (17.3)	0.125
Alcohol use, n (%)	1 (1.9)	2 (3.8)	1.000
Fell last year, n (%)	21 (40.4)	8 (15.4)	**0.009**
Living status, n (%)			
Family	39 (75.0)	37 (71.2)	0.825
Alone	13 (25.0)	15 (28.8)	
Social activities, n (%)			
Yes	22 (42.3)	30 (57.7)	0.170
No	30 (57.7)	22 (42.3)	
Disease duration (years)	7.97±5.61	—	—
Hoehn and Yahr stage n (%)		—	—
Stage 1	1 (1.9)	—	
Stage 1.5	2 (3.8)	—	
Stage 2	8 (15.4)	—	
Stage 2.5	26 (50.0)	—	
Stage 3	15 (28.8)	—	
LEDD (mg/day)	631.35±356.19	—	—

BMI: Body mass index, LEDD: Levodopa equivalent daily dose, n: Number, %: Percentage. Continuous variables are presented as mean ± standard deviation and categorical variables as n (%)

Table 2. Comparison of sarcopenia risk and sarcopenia disease risk perception

	Parkinson's disease (n=52)	Controls (n=52)	p-value
MSRA-7 score	30.19±7.19	33.65±5.98	**0.028**
MSRA-7 ≤30, n (%)	24 (46.2)	13 (25.0)	**0.041**
SDRPS-TR susceptibility score	30.50±3.89	29.63±3.91	0.346
SDRPS-TR severity score	26.92±3.36	26.15±3.82	0.340
SDRPS-TR total score	57.42±7.11	55.79±7.55	0.287

MSRA-7: Mini sarcopenia risk assessment-7, SDRPS-TR: Sarcopenia disease risk perception scale-Turkish version, n: Number, %: Percentage. Continuous variables are presented as mean ± standard deviation and categorical variables as n (%)

p=0.041). No significant differences were observed between the groups regarding SDRPS-TR susceptibility, severity, or total scores (all p>0.05).

Patients with PD were further compared according to MSRA-7-defined sarcopenia risk (Table 3). Patients at high sarcopenia risk were significantly older than those at low risk (p=0.027). They also reported a significantly higher frequency of falls during the previous year (p=0.006) and were more likely to live alone (p=0.025). No significant differences were observed between the groups regarding sex, BMI, education level, disease duration, Hoehn and Yahr stage, LEDD, hypertension, type 2 diabetes mellitus, dyslipidemia, coronary artery disease, polypharmacy, current smoking, alcohol use, social activities, or SDRPS-TR susceptibility, severity, and total scores (all p>0.05).

The correlations between clinical variables and SDRPS-TR scores are presented in Table 4. Age showed significant negative correlations with the SDRPS-TR susceptibility (r=-0.366, p=0.008), severity (r=-0.405, p=0.003), and total scores (r=-0.398, p=0.003). Education level showed significant positive correlations with the SDRPS-TR susceptibility (r=0.312, p=0.024), severity (r=0.346, p=0.012), and total scores (r=0.352, p=0.011). No significant correlations were observed between SDRPS-TR scores and BMI, disease duration, Hoehn and Yahr stage, LEDD, or MSRA-7 score (all p>0.05).

As shown in Table 5, participants living with their families had significantly higher SDRPS-TR susceptibility (p=0.001), severity (p=0.002), and total (p=0.001) scores than those living alone. Similarly, participants who reported engaging in social activities had significantly higher SDRPS-TR susceptibility (p=0.010), severity (p=0.006), and total (p=0.006) scores than those without social activities.

Discussion

The present study evaluated sarcopenia risk and SDRPS in patients with PD. Compared with healthy controls, patients with PD had significantly lower MSRA-7 scores and a higher prevalence of high sarcopenia risk. However, SDRPS did not differ significantly between the two groups. Among patients with PD, high sarcopenia risk was associated with older age, a history of falls, and living alone. In addition, increasing age was associated with lower sarcopenia risk perception, whereas higher educational level was positively correlated with SDRPS-TR scores. Moreover, participants living with their families and those engaging in social activities had significantly higher SDRPS-TR scores.

Sarcopenia is a muscle disease characterized by the progressive loss of muscle strength, muscle mass, and physical performance and is associated with an increased risk of falls, disability, hospitalization, and mortality (22,23). Growing evidence suggests that sarcopenia is highly prevalent in patients with PD and contributes to poorer functional outcomes and disease progression (24,25). Although the underlying mechanisms have not been fully elucidated, physical inactivity, malnutrition, chronic inflammation, mitochondrial dysfunction, oxidative stress, and neuromuscular dysfunction have been proposed as potential mechanisms contributing to the development of sarcopenia in patients with PD (25). Consistent with these findings, patients with PD in the present study had significantly lower MSRA-7 scores than healthy controls. High sarcopenia risk was identified in 46.2% of patients with PD compared with 25.0% of healthy controls, representing a statistically significant difference. The prevalence of sarcopenia in the general older population has been reported to range from approximately 6% to 22% and increases substantially with advancing age (26,27). In contrast, the reported prevalence of sarcopenia in patients with PD varies considerably according to the diagnostic criteria and assessment methods used, ranging from approximately 17% to 58% (5,28,29).

Within the PD group, patients at high sarcopenia risk were significantly older, had a higher prevalence of falls during the previous year, and were more likely to live alone than

Table 3. Comparison of patients with Parkinson's disease according to sarcopenia risk (MSRA-7)

	MSRA-7 ≤30 (n=24)	MSRA-7 >30 (n=28)	p-value
Age (years)	76.42±7.21	71.93±5.78	**0.027**
Gender, n (%)			
Female	11 (45.8)	14 (50.0)	0.983
Male	13 (54.2)	14 (50.0)	
BMI (kg/m²)	26.56±3.66	25.91±4.62	0.532
Education level, n (%)			
Illiterate	8 (33.3)	2 (7.1)	0.073
Primary school	4 (16.7)	9 (32.1)	
Middle school	1 (4.2)	4 (14.3)	
High school	5 (20.8)	3 (10.7)	
University	6 (25.0)	10 (35.7)	
Disease duration (years)	7.64±4.96	8.26±6.19	0.826
Hoehn and Yahr stage, n (%)			
Stage 1	0 (0.0)	1 (3.6)	0.867
Stage 1.5	1 (4.2)	1 (3.6)	
Stage 2	3 (12.5)	5 (17.9)	
Stage 2.5	13 (54.2)	13 (46.4)	
Stage 3	7 (29.2)	8 (28.6)	
LEDD (mg/day)	693.75±366.56	577.86±344.62	0.263
Hypertension, n (%)	8 (33.3)	13 (46.4)	0.499
Type 2 diabetes mellitus, n (%)	3 (12.5)	4 (14.3)	1.000
Dyslipidemia, n (%)	6 (25.0)	9 (32.1)	0.795
Coronary artery disease, n (%)	0 (0.0)	2 (7.1)	0.493
Polypharmacy (>5 medications), n (%)	11 (45.8)	10 (35.7)	0.64
Current smoking, n (%)	2 (8.3)	1 (3.6)	0.590
Alcohol use, n (%)	0 (0.0)	1 (3.6)	1.000
Fell last year, n (%)	15 (62.5)	6 (21.4)	**0.006**
Living status, n (%)			
Family	14 (58.3)	25 (89.3)	**0.025**
Alone	10 (41.7)	3 (10.7)	
Social activities, n (%)			
Yes	7 (29.2)	15 (53.6)	0.135
No	17 (70.8)	13 (46.4)	
SDRPS-TR susceptibility score	29.88±4.62	31.04±3.13	0.471
SDRPS-TR severity score	26.29±3.62	27.46±3.08	0.356
SDRPS-TR total score	56.17±8.17	58.50±6.00	0.368

BMI: Body mass index, LEDD: Levodopa equivalent daily dose, MSRA-7: Mini sarcopenia risk assessment-7, SDRPS-TR: Sarcopenia disease risk perception scale-Turkish version, n: Number, %: Percentage. Continuous variables are presented as mean ± standard deviation and categorical variables as n (%)

those at low risk. Advanced age is recognized as one of the strongest risk factors for sarcopenia because aging is accompanied by progressive declines in muscle mass, muscle strength, neuromuscular function, and regenerative capacity (30). A bidirectional relationship between sarcopenia and falls can be considered. Reduced muscle strength and impaired physical performance increase the risk of falls, whereas falls may further exacerbate physical inactivity, functional decline, and subsequent muscle loss, thereby accelerating the progression of sarcopenia (31). Regarding the association between living alone and sarcopenia, individuals living alone are more likely to experience social isolation, which may contribute to the development of sarcopenia through behavioral and

Table 4. Correlations between clinical variables and sarcopenia disease risk perception in Parkinson's disease (n=52)

	SDRPS-TR susceptibility		SDRPS-TR severity		SDRPS-TR total	
	r	p	r	p	r	p
Age (years)	-0.366	**0.008**	-0.405	**0.003**	-0.398	**0.003**
BMI (kg/m^2)	-0.186	0.186	-0.142	0.315	-0.153	0.279
Disease duration (years)	0.169	0.232	0.192	0.172	0.182	0.198
Education level	0.312	**0.024**	0.346	**0.012**	0.352	**0.011**
Hoehn and Yahr stage	0.008	0.954	-0.041	0.775	-0.028	0.844
LEDD (mg/day)	0.237	0.091	0.266	0.056	0.260	0.063
MSRA-7 score	-0.014	0.921	-0.003	0.981	0.001	0.997

BMI: Body mass index, LEDD: Levodopa equivalent daily dose, MSRA-7: Mini sarcopenia risk assessment-7, SDRPS-TR: Sarcopenia disease risk perception scale-Turkish version, r: Spearman's correlation coefficient

biological mechanisms, including reduced physical activity, inadequate nutritional intake, chronic inflammation, neuroendocrine dysregulation, and oxidative stress (32).

An important finding of the present study was that, despite the significantly higher risk of sarcopenia among patients with PD, perceived susceptibility, perceived severity, and total SDRPS scores did not differ significantly between patients with PD and healthy controls. This finding suggests that the increased risk of sarcopenia in PD is not accompanied by a parallel increase in disease awareness or perceived susceptibility and severity. Given that risk perception is an important determinant of health-related behaviors, insufficient awareness of sarcopenia may represent a barrier to the adoption of preventive strategies, including regular exercise, adequate protein intake, and early screening (14). In addition, increasing age was associated with lower sarcopenia risk perception, whereas higher educational level, living with family, and participation in social activities were associated with higher SDRPS-TR scores.

Studies evaluating SDRPS remain scarce. Amin et al. (33) investigated the relationships among fall efficacy, sarcopenia risk perception, and health behaviors in community-dwelling older adults and found that greater fall efficacy was associated with lower sarcopenia risk perception, with health behaviors partially mediating this association. The authors interpreted these findings in accordance with Bandura's self-efficacy theory, suggesting that individuals who perceive greater control over a given situation tend to perceive health threats as less severe. In another study, Guo et al. (34) assessed sarcopenia awareness among community-dwelling older adults and reported that only 17.8% of participants had previously heard of sarcopenia. They further demonstrated that older age, lower educational level, living alone or with children, and lack of participation in social activities were associated with lower levels of sarcopenia knowledge. Similarly, Sun et al. (35) conducted a cross-sectional survey to evaluate knowledge, attitudes, and practices regarding sarcopenia among adults with type 2 diabetes. They reported that participants had limited knowledge despite generally favorable attitudes and moderate practice behaviors related to sarcopenia. Furthermore, higher knowledge levels were positively associated with more favorable attitudes and better health-related practices, suggesting that belonging to a high-risk population alone does not necessarily translate into greater awareness of sarcopenia. These findings are largely consistent with our results and support the concept that sarcopenia awareness and disease risk perception are influenced not only by objective clinical risk but also by educational, social, and behavioral factors. Therefore, improving awareness of sarcopenia through patient education, routine screening, and individualized counseling should be considered an integral component of the clinical management of patients with PD, particularly those at increased risk of sarcopenia, to facilitate the adoption of preventive health behaviors and promote early identification and management of the condition.

Study Limitations

This study has several limitations. First, its cross-sectional design precludes any inference of causal relationships between sarcopenia risk, sarcopenia risk perception, and the associated clinical or sociodemographic factors. Second, the study was conducted at a single center with a relatively small sample size, which may limit the generalizability of the findings. Third, sarcopenia risk was assessed using the MSRA-7 questionnaire rather than a diagnostic evaluation based on established consensus criteria, such as the European Working Group on Sarcopenia in Older People 2 recommendations; therefore, the results reflect sarcopenia

Table 5. Comparison of SDRPS-TR scores according to categorical variables in Parkinson's disease (n=52)

	SDRPS-TR susceptibility		SDRPS-TR severity		SDRPS-TR total	
	Mean ± SD	**p**	**Mean ± SD**	**p**	**Mean ± SD**	**p**
Gender						
Male	30.37±3.31	0.919	26.85±3.17	0.890	57.22±6.26	0.971
Female	30.64±4.51		27.00±3.62		57.64±8.06	
Hypertension						
Yes	30.71±4.60	0.858	27.14±3.58	1.000	57.86±8.11	0.985
No	30.35±3.41		26.77±3.25		57.13±6.47	
Type 2 diabetes mellitus						
Yes	29.00±3.06	0.246	25.43±2.70	0.186	54.43±5.71	0.202
No	30.73±3.99		27.16±3.42		57.89±7.25	
Dyslipidemia						
Yes	30.00±5.17	0.976	26.53±3.93	0.879	56.53±9.01	0.879
No	30.70±3.31		27.08±3.15		57.78±6.29	
Coronary artery disease						
Yes	29.50±2.12	0.615	25.50±2.12	0.416	55.00±4.24	0.535
No	30.54±3.95		26.98±3.40		57.52±7.21	
Polypharmacy (>5 medications)						
Yes	30.57±2.87	0.948	26.81±2.58	0.851	57.38±5.43	0.881
No	30.45±4.50		27.00±3.84		57.45±8.14	
Current smoking						
Yes	31.00±2.65	0.843	27.33±1.53	0.828	58.33±4.16	0.829
No	30.47±3.97		26.90±3.45		57.37±7.28	
Alcohol use[†]						
Yes	30.00	—	27.00	—	57.00	—
No	30.51±3.93		26.92±3.39		57.43±7.18	
Fell last year						
Yes	29.62±4.92	0.302	26.05±3.89	0.209	55.67±8.75	0.243
No	31.10±2.95		27.52±2.86		58.61±5.60	
Living status						
Family	31.64±2.96	**0.001**	27.85±2.79	**0.002**	59.49±5.57	**0.001**
Alone	27.08±4.44		24.15±3.51		51.23±7.83	
Social activities						
Yes	32.32±3.43	**0.010**	28.55±3.16	**0.006**	60.86±6.32	**0.006**
No	29.17±3.71		25.73±3.03		54.90±6.67	

SDRPS-TR: Sarcopenia disease risk perception scale-Turkish version, SD: Standard deviation, [†]: Statistical comparison was not performed because only one participant reported alcohol use

risk rather than confirmed sarcopenia. Finally, although the SDRPS-TR provides valuable information regarding individuals' perceptions of sarcopenia, it is a self-reported instrument and may be subject to recall and response bias.

Conclusion

In conclusion, sarcopenia is an increasingly important clinical condition with a high prevalence among patients with PD. However, our findings indicate that the increased risk of sarcopenia in patients with PD is not accompanied by a higher level of sarcopenia risk perception. These findings highlight the need for strategies aimed at improving awareness of sarcopenia. Therefore, integrating routine sarcopenia screening with patient education and awareness-promoting interventions may facilitate preventive health behaviors and contribute to the early identification and management of sarcopenia in patients with PD.

Ethics

Ethics Committee Approval: The study protocol was approved by the Local Institutional Non-Interventional Clinical Research Ethics Committee of İstanbul Medipol University (decision no: 866, date: 14.05.2026).

Informed Consent: All participants provided written informed consent before enrollment.

Footnotes

Authorship Contributions

Surgical and Medical Practices: H.B.M., A.T.P., M.Y., E.T., E.A., S.T., Concept: H.B.M., A.E., A.T.P., M.Y., E.T., E.A., S.T., Design: H.B.M., A.E., A.T.P., M.Y., E.T., E.A., S.T., Data Collection or Processing: H.B.M., E.T., S.T., Analysis or Interpretation: H.B.M., A.E., A.T.P., M.Y., E.T., E.A., S.T., Literature Search: H.B.M., A.E., A.T.P., M.Y., E.T., S.T., Writing: H.B.M., A.E., A.T.P., M.Y., E.T., E.A., S.T.

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

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

References

1. Costa HN, Esteves AR, Empadinhas N, Cardoso SM. Parkinson's disease: a multisystem disorder. Neurosci Bull. 2023;39(1):113-124.
2. GBD 2016 Parkinson's Disease Collaborators. Global, regional, and national burden of Parkinson's disease, 1990-2016: a systematic analysis for the global burden of disease study 2016. Lancet Neurol. 2018;17(11):939-953.
3. Leite Silva ABR, Gonçalves de Oliveira RW, Diógenes GP, de Castro Aguiar MF, Sallem CC, Lima MPP, et al. Premotor, nonmotor and motor symptoms of Parkinson's disease: a new clinical state of the art. Ageing Res Rev. 2023;84:101834.
4. Yazar T, Yazar HO, Zayimoğlu E, Çankaya S. Incidence of sarcopenia and dynapenia according to stage in patients with idiopathic Parkinson's disease. Neurol Sci. 2018;39(8):1415-1421.
5. Peball M, Mahlknecht P, Werkmann M, Marini K, Murr F, Herzmann H, et al. Prevalence and associated factors of sarcopenia and frailty in Parkinson's disease: a cross-sectional study. Gerontology. 2019;65(3):216-228.
6. Vetrano DL, Pisciotta MS, Laudisio A, Lo Monaco MR, Onder G, Brandi V, et al. Sarcopenia in Parkinson disease: comparison of different criteria and association with disease severity. J Am Med Dir Assoc. 2018;19(6):523-527.
7. Cai Y, Feng F, Wei Q, Jiang Z, Ou R, Shang H. Sarcopenia in patients with Parkinson's disease: a systematic review and meta-analysis. Front Neurol. 2021;12:598035.
8. Marzetti E, Calvani R, Tosato M, Cesari M, Di Bari M, Cherubini A, et al. Sarcopenia: an overview. Aging Clin Exp Res. 2017;29(1):11-17.
9. Chen LK, Woo J, Assantachai P, Auyeung TW, Chou MY, Iijima K, et al. Asian Working Group for Sarcopenia: 2019 consensus update on sarcopenia diagnosis and treatment. J Am Med Dir Assoc. 2020;21(3):300-307.e2.
10. Sobestiansky S, Michaelsson K, Cederholm T. Sarcopenia prevalence and associations with mortality and hospitalisation by various sarcopenia definitions in 85-89 year old community-dwelling men: a report from the ULSAM study. BMC Geriatr. 2019;19(1):318.
11. Kamaci Sener D, Deniz O, Gemci E, Gonullu S, Kamisli S. The association between probable sarcopenia, non-motor symptoms, and fall-related outcomes in Parkinson's disease. Sci Rep. 2026;16(1):5650.
12. Gomes de Luna JR, Lima DP, Gomes VC, Lopes FKM, Brito de Almeida S, Barroso PLS, et al. Sarcopenia, body composition, physical performance, and clinical features in Parkinson's disease: a cross-sectional comparative study. Front Endocrinol (Lausanne). 2026;17:1793254.
13. Almeida SB, Lima DP, Luna JRG, Viana Júnior AB, Roriz-Filho JS, Alencar ÁP, et al. Physical performance and clinical correlates of sarcopenia in Parkinson disease: a cross-sectional study. Arq Neuropsiquiatr. 2026;84(2):1-11.
14. Ferrer R, Klein WM. Risk perceptions and health behavior. Curr Opin Psychol. 2015;5:85-89.
15. Meral HB, Tolu S, Noyan Z. Validation and cross-cultural adaptation of the Sarcopenia Disease Risk Perception Scale among Turkish older adults. Turk J Phys Med Rehabil. 2026;72(2):197-206.
16. Postuma RB, Berg D, Stern M, Poewe W, Olanow CW, Oertel W, et al. MDS clinical diagnostic criteria for Parkinson's disease. Mov Disord. 2015;30(12):1591-1601.
17. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12(3):189-198.
18. Tomlinson CL, Stowe R, Patel S, Rick C, Gray R, Clarke CE. Systematic review of levodopa dose equivalency reporting in Parkinson's disease. Mov Disord. 2010;25(15):2649-2653.
19. Rossi AP, Micciolo R, Rubele S, Fantin F, Caliari C, Zoico E, et al. Assessing the risk of sarcopenia in the elderly: the mini sarcopenia risk assessment (MSRA) questionnaire. J Nutr Health Aging. 2017;21(6):743-749.
20. Yilmaz O, Kilic C, Seker N, Şentürk Durmuş N, Kucukdagli P, Ozulu Turkmen B, et al. Assessing the applicability of the mini sarcopenia risk assessment questionnaire in Turkish older adult population: a validation study. Acta Clin Belg. 2026;81(2):84-94.
21. Zhang W, Sun Z, Wang J, Wu Y. Development and validation of the sarcopenia disease risk perception scale for older adults. BMC Geriatr. 2024;24(1):876.
22. Park WT, Shon OJ, Kim GB. Multidisciplinary approach to sarcopenia: a narrative review. J Yeungnam Med Sci. 2023;40(4):352-363.
23. Sayer AA, Cruz-Jentoft A. Sarcopenia definition, diagnosis and treatment: consensus is growing. Age Ageing. 2022;51(10):afac220.
24. Morimoto R, Kitagawa K, Todo K, Iijima M. Clinical analysis of sarcopenia prevalence and its influencing factors in patients with Parkinson's disease. Front Aging Neurosci. 2025;17:1718723.
25. Gui M, Lv L, Hu S, Qin L, Wang C. Sarcopenia in Parkinson's disease: from pathogenesis to interventions. Metabolism. 2025;169:156272.

26. Bianchi L, Abete P, Bellelli G, Bo M, Cherubini A, Corica F, et al. Prevalence and clinical correlates of sarcopenia, identified according to the EWGSOP definition and diagnostic algorithm, in hospitalized older people: the GLISTEN study. J Gerontol A Biol Sci Med Sci. 2017;72(11):1575-1581.

27. Morley JE, Anker SD, von Haehling S. Prevalence, incidence, and clinical impact of sarcopenia: facts, numbers, and epidemiology-update 2014. J Cachexia Sarcopenia Muscle. 2014;5(4):253-259.

28. Tan AH, Hew YC, Lim SY, Ramli NM, Kamaruzzaman SB, Tan MP, et al. Altered body composition, sarcopenia, frailty, and their clinico-biological correlates, in Parkinson's disease. Parkinsonism Relat Disord. 2018;56:58-64.

29. Ozer FF, Akın S, Gultekin M, Zararsız GE. Sarcopenia, dynapenia, and body composition in Parkinson's disease: are they good predictors of disability? A case-control study. Neurol Sci. 2020;41(2):313-320.

30. Araújo LP, Figueiredo Godoy AC, Fortes Frota F, Barbalho Lamas C, Quesada K, Rucco Penteado Detregiachi C, et al. Sarcopenia in the aging process: pathophysiological mechanisms, clinical implications, and emerging therapeutic approaches. Int J Mol Sci. 2025;26(24):12147.

31. Yeung SSY, Reijnierse EM, Pham VK, Trappenburg MC, Lim WK, Meskers CGM, et al. Sarcopenia and its association with falls and fractures in older adults: a systematic review and meta-analysis. J Cachexia Sarcopenia Muscle. 2019;10(3):485-500.

32. Yang J, Huang J, Yang X, Li S, Wu X, Ma X. The association of living alone and social isolation with sarcopenia: a systematic review and meta-analysis. Ageing Res Rev. 2023;91:102043.

33. Amin SM, Alkubati SA, Hra M, Zoromba MA, Abdallah HMM, Ghoniem NGE, et al. Perceiving the invisible: the mediating role of health behaviour in the link between fall efficacy and sarcopenia risk perception among community-dwelling older adults. Geriatr Nurs. 2026;71:104082.

34. Guo W, Wang T, Dai H, Wang Y, Zhou Z, Wang Y, et al. The knowledge level and influencing factors of sarcopenia among Chinese community-dwelling older adults. PLoS One. 2025;20(10):e0333557.

35. Sun X, Liu W, Han S, Wang M, Liu Z, Jiang W, et al. Knowledge, attitudes, and practices about sarcopenia in adults with type 2 diabetes: a cross-sectional survey. J Diabetes Res. 2026;2026(1):e3057863.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(3):414-426
DOI: 10.4274/BMB.galenos.2026.97442

Analysis of Patients Who Underwent Surgery for Gastric Gastrointestinal Stromal Tumors

Mide Gastrointestinal Stromal Tümör Nedeniyle Opere Edilen Hastaların Analizi

Nurettin Şahin[1], Turgut Dönmez[1], Mahmut Said Değerli[1], Ferman Tevfik Özyalvaç[1], Süleyman Büyükaşık[2], Allanur Mametkulyyev[1], Duhan Kanmaz[1], Ahmet Sürek[1]

[1]University of Health Sciences Turkey, İstanbul Bakırköy Dr. Sadi Konuk Training and Research Hospital, Department of General Surgery, İstanbul, Turkey
[2]İstanbul Aydın University Faculty of Medicine, Department of General Surgery, İstanbul, Turkey

Abstract

Objective: To evaluate the clinicopathological characteristics, surgical management, and medium-term outcomes of patients who underwent surgery for gastric gastrointestinal stromal tumors (GISTs).

Method: This single-center retrospective cohort study included 39 consecutive patients who underwent surgery for histopathologically confirmed gastric GIST between January 2016 and June 2026. Clinical, surgical, pathological, and follow-up data were reviewed. Overall survival was estimated using the Kaplan-Meier method, and median follow-up was calculated using the reverse Kaplan-Meier method.

Results: The mean age was 60.4±12.6 years. Laparoscopic surgery was performed in 30 patients (76.9%), and wedge resection in 34 (87.2%). The median tumor diameter was 4.5 cm. Among 34 patients with an available risk classification, 28 (82.4%) had very low- or low-risk tumors, four had intermediate-risk tumors, and two had high-risk tumors. R0 resection was achieved in all 38 patients with evaluable margin status. Six patients (15.4%) experienced Clavien-Dindo grade II complications. No documented recurrence was identified among the 38 patients with evaluable recurrence status. Median follow-up calculated using the reverse Kaplan-Meier method was 42 months. Overall survival was 97.4% at 24 and 48 months and 86.6% at 60 months; however, only nine patients remained at risk at 60 months.

Conclusion: Organ-preserving wedge resection and a laparoscopic approach were frequently used in selected patients with localized gastric GIST, and R0 resection was achieved in all patients with

Öz

Amaç: Cerrahi uygulanan gastrik gastrointestinal stromal tümörlerin (GIST) klinikopatolojik özelliklerini, cerrahi yönetimini ve orta dönem sonuçlarını değerlendirmektir.

Yöntem: Bu tek merkezli retrospektif kohort çalışmasına, Ocak 2016-Haziran 2026 tarihleri arasında histopatolojik olarak doğrulanmış gastrik GIST nedeniyle cerrahi uygulanan 39 ardışık hasta dahil edildi. Klinik, cerrahi, patolojik ve takip verileri incelendi. Genel sağkalım Kaplan-Meier, medyan takip süresi ise ters Kaplan-Meier yöntemiyle hesaplandı.

Bulgular: Hastaların ortalama yaşı 60,4±12,6 yıldı. Otuz hastaya (%76,9) laparoskopik cerrahi, 34 hastaya (%87,2) wedge rezeksiyon uygulandı. Medyan tümör çapı 4,5 cm idi. Risk sınıfı bilinen 34 hastanın 28'i (%82,4) çok düşük veya düşük risk grubundaydı; dört hasta orta, iki hasta yüksek riskliydi. Cerrahi sınır durumu değerlendirilebilen 38 hastanın tamamında R0 rezeksiyon sağlandı. Altı hastada (%15,4) Clavien-Dindo derece II komplikasyon gelişti. Nüks durumu değerlendirilebilen 38 hastanın hiçbirinde belgelenmiş nüks saptanmadı. Ters Kaplan-Meier yöntemine göre medyan takip süresi 42 aydı. Genel sağkalım 24. ve 48. aylarda %97,4, 60. ayda %86,6 olarak hesaplandı; ancak 60. ayda risk altında yalnızca dokuz hasta bulunmaktaydı.

Sonuç: Seçilmiş lokalize gastrik GIST hastalarında organ koruyucu wedge rezeksiyon ve laparoskopik yaklaşım sık uygulanmış, değerlendirilebilir hastaların tamamında R0 rezeksiyon sağlanmıştır. Orta dönem sonuçlar olumlu olmakla birlikte düşük

Address for Correspondence: Nurettin Şahin, University of Health Sciences Turkey, İstanbul Bakırköy Dr. Sadi Konuk Training and Research Hospital, Department of General Surgery, İstanbul, Turkey
E-mail: dr.nurettinsahin@gmail.com **ORCID:** orcid.org/0000-0001-7440-0574
Received: 03.08.2026 **Accepted:** 04.09.2026 **Publication Date:** 29.09.2026

Cite this article as: Şahin N, Dönmez T, Değerli MS, Özyalvaç FT, Büyükaşık S, Mametkulyyev A, et al. Analysis of patients who underwent surgery for gastric gastrointestinal stromal tumors. Bagcilar Med Bull. 2026;11(3):414-426

Abstract

evaluable margin status. Although medium-term outcomes were favorable, the predominance of low-risk tumors, limited sample size, and declining number of patients at risk during later follow-up preclude definitive conclusions regarding long-term oncological outcomes.

Keywords: Gastric gastrointestinal stromal tumor, minimally invasive surgery, wedge resection

Öz

riskli tümörlerin ağırlığı, sınırlı örneklem büyüklüğü ve ileri takip dönemlerindeki azalan hasta sayısı uzun dönem onkolojik çıkarımları sınırlandırmaktadır.

Anahtar kelimeler: Gastrik gastrointestinal stromal tümör, minimal invaziv cerrahi, wedge rezeksiyon

Introduction

Gastrointestinal stromal tumors (GISTs) are the most common mesenchymal neoplasms of the gastrointestinal tract, with the stomach representing their predominant site of origin. Their biological behavior is heterogeneous: Small tumors with low mitotic activity may remain clinically silent for prolonged periods, whereas more aggressive lesions may exhibit local invasion, peritoneal dissemination, or distant metastasis. The identification of oncogenic driver alterations, particularly those involving KIT and platelet-derived growth factor receptor alpha (PDGFRA), has substantially transformed the diagnostic classification and systemic treatment of GISTs (1-6).

The clinical presentation of gastric GISTs varies according to tumor size, anatomical location, growth pattern, and mucosal ulceration. Small lesions may be detected incidentally, whereas larger tumors may present with abdominal pain, gastrointestinal bleeding, early satiety, anemia, or an abdominal mass. Contrast-enhanced computed tomography is the principal modality for assessing tumor size, adjacent structures, and metastatic disease. Endoscopic ultrasonography and tissue sampling are used selectively when the diagnosis is uncertain, neoadjuvant treatment is considered, or pathological confirmation may alter management. Diagnosis is based on morphology and immunohistochemical markers, particularly CD117 and DOG1, while molecular testing refines classification and informs prognostic, predictive, and treatment decisions (1-4,6).

Surgery is the mainstay of treatment for localized and resectable gastric GISTs. The objective is macroscopically complete resection with a microscopically negative margin while preserving the tumor pseudocapsule and avoiding rupture. Routine lymphadenectomy is not recommended because lymph node metastasis is uncommon. The stomach often permits organ-preserving wedge resection instead of more extensive resection. A laparoscopic approach may be considered for appropriately sized and located tumors without adjacent organ invasion and has been associated with shorter hospital stay and acceptable perioperative outcomes in selected patients. The approach should nevertheless be individualized according to tumor anatomy and technical complexity (1-4,7,8).

Tumor size, mitotic activity, and tumor rupture are among the principal determinants of recurrence risk in gastric GISTs (1-6). In retrospective series with heterogeneous follow-up durations, oncological outcomes should be assessed using time-to-event methods that appropriately account for censored observations. Although single-center series can reflect routine surgical practice and real-world outcomes, missing data and the duration of follow-up must be reported transparently. The primary aim of this study was to evaluate the clinical, surgical, histopathological, and immunohistochemical characteristics of patients who underwent surgery for histopathologically confirmed gastric GIST. The secondary aim was to assess postoperative follow-up findings, documented recurrences, and medium-term survival outcomes using appropriate time-to-event methods.

Materials and Methods

Study Design and Patient Selection

This single-center, retrospective observational cohort study was conducted at the department of general surgery of University of Health Sciences Turkey, İstanbul Bakırköy Dr. Sadi Konuk Training and Research Hospital between January 2016 and June 2026. The study was reported in accordance with the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations (9).

Adult patients who underwent surgical resection for a presumed gastric subepithelial tumor during the study period were assessed for eligibility. Consecutive patients with a definitive histopathological diagnosis of gastric GIST and available essential clinical, operative, and histopathological

data were included. Patients with insufficient essential records and those whose final histopathological diagnosis was other than gastric GIST were excluded.

A total of 64 potentially eligible patients were screened during the study period. Four patients with incomplete essential records and 21 patients with a final histopathological diagnosis other than gastric GIST were excluded. The final study cohort comprised 39 patients with histopathologically confirmed gastric GIST. The patient selection process is presented in Figure 1.

The study was approved by the Local Clinical Research Ethics Committee of University of Health Sciences Turkey, İstanbul Bakırköy Dr. Sadi Konuk Training and Research Hospital before its initiation (approval date: June 19, 2026; decision no: 2026-14-26). Written informed consent for the use of clinical data for research purposes was obtained from all patients before their inclusion in the study. The study was conducted in accordance with the principles of the Declaration of Helsinki.

Data Sources and Clinical Variables

Data were retrospectively obtained by reviewing the hospital electronic medical records, operative notes, imaging reports, final pathology reports, and outpatient follow-up records. The recorded variables included age, sex, body mass index (BMI), American Society of Anesthesiologists (ASA) physical status classification, Charlson comorbidity index (CCI), history of previous abdominal surgery, smoking and alcohol consumption, and presenting symptoms.

The preoperative assessment included whether upper gastrointestinal endoscopy, endoscopic ultrasonography, abdominal ultrasonography, contrast-enhanced computed tomography, magnetic resonance imaging, and positron emission tomography had been performed, together with the reported findings of these examinations. Preoperative tissue sampling was performed under endoscopic ultrasonographic guidance using fine-needle aspiration. The histopathological results were categorized as compatible with GIST or non-diagnostic. Core tissue acquisition was not performed. The presence of distant metastasis at diagnosis was determined from imaging findings and clinical records.

Surgical Assessment

The surgical variables comprised the anatomical location of the tumor within the stomach, its relationship with the gastric curvature, the surgical approach, the type of gastric resection, and operative duration. Tumor location was categorized as the antrum, body, or fundus. Its relationship with the gastric curvature was classified as greater

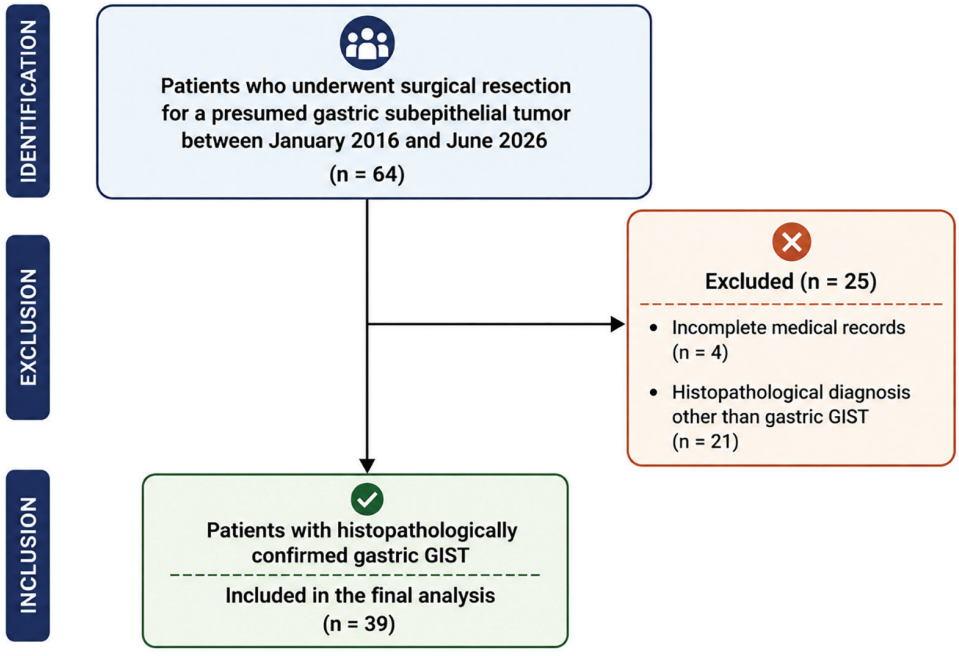

Figure 1. Flow diagram of patient selection
GIST: Gastrointestinal stromal tumor

curvature, lesser curvature, or not reported. The surgical approach was recorded as laparoscopic or open, whereas the type of resection was categorized as wedge resection, subtotal gastrectomy, or total gastrectomy.

The surgical approach was not selected according to a standardized protocol predefined for the purposes of this study. The treatment team determined the operative approach based on tumor size and anatomical location, its relationship with adjacent structures, technical resectability, and the need for concomitant surgical procedures.

Postoperative Assessment

Postoperative length of hospital stay was calculated from the date of surgery to the date of discharge. Postoperative complications were graded according to the Clavien-Dindo classification. Ninety-day mortality was defined as death from any cause within the first 90 days after surgery. Adjuvant imatinib use and all-cause mortality during follow-up were also recorded.

Histopathological Assessment

Histopathological data were obtained from the final pathology reports. The assessed variables included histological cell type, cellularity, cellular atypia, maximum tumor diameter, mitotic count per 5 mm², tumor necrosis, lymphovascular invasion, perineural invasion, resection margin status, integrity of the resection specimen, and the risk category documented in the pathology report. Cellularity and cellular atypia were recorded as reported in the original pathology reports, and no retrospective classification was performed for missing data.

Risk categories were recorded as very low, low, intermediate, or high according to the modified National Institutes of Health classification documented in the pathology reports (10). Patients without a reported risk category were not retrospectively assigned to a new risk group.

A microscopically tumor-free resection margin was considered an R0 resection. Patients with unavailable resection margin information were not classified as having an R1 resection. Resection margin status was presented in the tables as negative or not reported, and the R0 resection rate among evaluable patients was calculated separately.

Cases described as having "incomplete specimen integrity" in the pathology report were identified separately. For the retrospective assessment of tumor rupture, tumor fracture or spillage into the peritoneal cavity, blood-stained ascites, gastrointestinal perforation at the tumor site, microscopic invasion of an adjacent organ, intralesional dissection or piecemeal resection, and incisional biopsy were considered criteria for tumor rupture (11). Tumor rupture was considered confirmed only when at least one predefined rupture criterion was documented in the operative or detailed pathology records. Therefore, incomplete specimen integrity alone was not regarded as definitive tumor rupture.

Immunohistochemical and Molecular Assessment

The recorded immunohistochemical findings included staining results for CD117, DOG1, CD34, smooth muscle actin (SMA), desmin, S100, and CD31. The result for each marker was classified as positive, negative, or not reported. The Ki-67 proliferation index was assessed as a continuous percentage and was also categorized as <10% or ≥10%. Patients without an available Ki-67 result were excluded from these assessments.

Molecular mutation analysis, including testing for KIT and PDGFRA mutations, had not been performed in any patient. Consequently, the relationships between mutation status and risk classification, adjuvant imatinib use, or oncological outcomes could not be evaluated.

Follow-up and Oncological Outcomes

Follow-up data were obtained from outpatient examinations, imaging reports, and hospital electronic records. Follow-up data were censored in June 2026. Follow-up duration was calculated from the date of surgery to the date of the last clinical contact or death.

Overall survival was defined as the interval from surgery to death from any cause. Patients who were alive at the end of follow-up were censored at the date of their last clinical contact. Recurrence was defined as local, peritoneal, or distant GIST recurrence documented by imaging, histopathological examination, or clinical records during follow-up. Patients with unavailable recurrence information were not considered recurrence-free and were excluded from the denominator used to calculate the recurrence rate.

Because no documented recurrence occurred during follow-up, separate Kaplan-Meier estimates of disease-free or recurrence-free survival were not generated. Recurrence outcomes were presented descriptively as numbers and percentages. Time-to-event analysis was performed only for overall survival.

Statistical Analysis

Continuous variables were summarized using the mean ± standard deviation, median, and minimum-maximum

values. Categorical variables were presented as numbers and percentages. The proportions of patients who underwent each preoperative examination were calculated using the entire study cohort, whereas percentages relating to examination findings were calculated using the number of patients who underwent the corresponding examination. Unless otherwise specified, percentages in the remaining tables were calculated using the entire study cohort, and unreported data were presented as a separate category. For proportions calculated among evaluable patients, the denominator was explicitly reported. No imputation was performed for missing data.

The study sample comprised all consecutive patients who met the eligibility criteria during the predefined study period, and no a priori sample size calculation was performed. Given the descriptive study design, limited sample size, non-randomized selection of the surgical approach, and unequal group sizes, no hypothesis-testing comparison was performed between the laparoscopic and open surgery groups. No multivariable model of prognostic factors was constructed.

Overall survival was estimated using the Kaplan-Meier product-limit method. The 95% confidence intervals for survival estimates were calculated using Greenwood's variance and a log-log transformation. The numbers of patients at risk at 0, 12, 24, 36, 48, 60, 84, and 120 months were displayed below the survival curve.

Median follow-up duration was estimated using the reverse Kaplan-Meier method, in which the death and censoring indicators were reversed (12). Observed follow-up duration was also summarized using the mean ± standard deviation, median, and minimum-maximum values. Statistical analyses were performed in Python version 3.12 using reproducible analysis code. Data management and descriptive analyses were conducted using pandas version 2.2.3 and SciPy version 1.17.0.

Results

Patient Selection and Baseline Clinical Characteristics

Of 64 patients assessed after surgical resection for a presumed gastric subepithelial tumor, four with incomplete essential records and 21 with a final diagnosis other than gastric GIST were excluded. The final cohort comprised 39 patients with histopathologically confirmed gastric GIST (Figure 1).

The mean age of the patients was 60.4±12.6 years, and the median age was 63 years (range, 22-78 years). The cohort included 22 women (56.4%) and 17 men (43.6%). Among the 37 patients with available BMI data, the mean BMI was 28.7±5.8 kg/m^2 and the median was 27.8 kg/m^2 (range, 18.5-46.7 kg/m^2). BMI data were unavailable for two patients. Seventeen patients (43.6%) were classified as ASA physical status II and 22 (56.4%) as ASA physical status III. The mean CCI was 2.7±1.5, with a median of 3 (range, 0-6). Nineteen patients (48.7%) had a history of previous abdominal surgery.

Abdominal pain was the most common presenting symptom (n=17, 43.6%). Gastrointestinal bleeding was recorded in seven patients (17.9%), dyspepsia in six (15.4%), and fatigue in two (5.1%). Seven patients (17.9%) were asymptomatic and were diagnosed incidentally during examinations performed for other reasons. Ten patients (25.6%) were smokers, and four (10.3%) reported alcohol consumption. The baseline clinical characteristics of the patients are presented in Table 1.

Preoperative Diagnostic Assessment

Upper gastrointestinal endoscopy and contrast-enhanced computed tomography were performed in all patients. Endoscopy demonstrated a subepithelial lesion considered compatible with GIST in 32 patients (82.1%), extrinsic compression in three (7.7%), and a mass that was not initially considered to represent a GIST in one (2.6%). No distinct endoscopic lesion was identified in three patients (7.7%).

Contrast-enhanced computed tomography demonstrated a mass compatible with GIST in 29 patients (74.4%), gastric wall thickening in four (10.3%), and a lesion initially considered to be of extragastric origin in two (5.1%). No finding suggestive of GIST was identified on computed tomography in four patients (10.3%).

Endoscopic ultrasonography was performed in 15 patients (38.5%), magnetic resonance imaging in 14 (35.9%), positron emission tomography in seven (17.9%), and abdominal ultrasonography in three (7.7%). Among the patients who underwent endoscopic ultrasonography, five had a submucosal lesion and five had a lesion originating from the muscularis propria; the layer of origin was not specified in the remaining five patients. Preoperative fine-needle aspiration biopsy (FNAB) was performed in 13 patients, yielding pathological findings compatible with GIST in 12 (92.3%); the sample obtained from one patient was non-diagnostic. No patient had distant metastasis at the time of diagnosis. The preoperative diagnostic findings are presented in Table 2.

Table 1. Baseline demographic and clinical characteristics of the study population (n=39)

		Min-max			Median	Mean ± SD/n-%		
Age (years)		22.0	-	78.0	63.0	60.4	±	12.6
Sex	Female					22		56.4
	Male					17		43.6
BMI (kg/m^2)		18.5	-	46.7	27.8	28.7	±	5.8
ASA score	II					17		43.6
	III					22		56.4
CCI		0.0	-	6.0	3.0	2.7	±	1.5
Previous abdominal surgery	Yes					19		48.7
	No					20		51.3
Presenting symptom	Abdominal pain					17		43.6
	Gastrointestinal bleeding					7		17.9
	Asymptomatic					7		17.9
	Dyspepsia					6		15.4
	Fatigue					2		5.1
Smoking status	Yes					10		25.6
	No					29		74.4
Alcohol consumption	Yes					4		10.3
	No					35		89.7

BMI: Body mass index, ASA: American Society of Anesthesiologists, CCI: Charlson comorbidity index, SD: Standard deviation, BMI data were available for 37 patients

Surgical Findings

The tumor was located in the fundus in 16 patients (41.0%), the body in 15 (38.5%), and the antrum in eight (20.5%). The relationship with the gastric curvature was reported in 33 patients: 20 tumors were associated with the lesser curvature and 13 with the greater curvature. This information was not reported in six patients.

A laparoscopic approach was used in 30 patients (76.9%), whereas nine patients (23.1%) underwent open surgery. Wedge resection was the most frequently performed procedure, accounting for 34 operations (87.2%). Subtotal gastrectomy was performed in three patients (7.7%) and total gastrectomy in two (5.1%).

The mean operative duration was 114.5 ± 85.1 minutes, and the median was 94 minutes (range, 45-515 minutes). The 515-minute operation included a concomitant oncological debulking procedure for ovarian cancer in addition to gastric resection. Surgical characteristics are presented in Table 3.

Histopathological Findings

The histological cell type was spindle-cell in 24 patients (61.5%), mixed in eight (20.5%), and epithelioid in five (12.8%). Cell type was not reported in two patients. Two separate GIST foci were identified in the gastric resection specimen of one patient.

The mean tumor diameter was 6.3±5.5 cm, and the median was 4.5 cm (range, 0.8-32 cm). The mean mitotic count was 3.0±2.8 per 5 mm^2, with a median of 2 per 5 mm^2 (range, 1-14 per 5 mm^2). The mitotic count was ≤5 per 5 mm^2 in 36 patients (92.3%) and >5 per 5 mm^2 in three patients (7.7%).

Tumor necrosis was present in 14 patients (35.9%) and absent in 23 (59.0%); necrosis status was not reported in two patients. According to the documented risk classification, 17 patients (43.6%) were categorized as very low risk, 11 (28.2%) as low risk, four (10.3%) as intermediate risk, and two (5.1%) as high risk. Risk category was not reported in five patients (12.8%).

R0 resection was achieved in all 38 patients with evaluable margin status; margin status was unavailable for one patient, who was not classified as having an R1 resection.

Specimen integrity was complete in 24 patients and incomplete in four; it was not reported in 11. Among 28 evaluable patients, the corresponding proportions were 85.7% and 14.3%. Lymphovascular and perineural invasion were evaluated in only five and three patients, respectively, and were absent in all evaluable cases. These variables were not interpreted further because of extensive missing data (Table 4).

Table 2. Preoperative evaluation of the patients (n=39)

		n	%
Upper gastrointestinal endoscopy		39	100.0%
Endoscopic findings	Findings compatible with GIST	32	82.1%
	Mass lesion (GIST not suspected)	1	2.6%
	Extrinsic compression	3	7.7%
	No lesion detected	3	7.7%
Endoscopic ultrasonography		15	38.5%
EUS findings	Submucosal lesion	5	33.3%
	Muscularis propria origin	5	33.3%
	Not specified	5	33.3%
Abdominal ultrasonography		3	7.7%
Ultrasonography findings	Mass detected	2	66.7%
	No mass detected	1	33.3%
Contrast-enhanced computed tomography		39	100.0%
Computed tomography findings	Findings compatible with GIST	29	74.4%
	Gastric wall thickening	4	10.3%
	Normal findings	4	10.3%
	Extragastric lesion	2	5.1%
Magnetic resonance imaging		14	35.9%
MRI findings	Findings compatible with GIST	10	71.4%
	Findings not suggestive of GIST	3	21.4%
	Extragastric lesion	1	7.1%
Positron emission tomography		7	17.9%
PET findings	Findings compatible with GIST	7	100.0%
Fine-needle aspiration biopsy		13	33.3%
Biopsy findings	Histopathological findings compatible with GIST	12	92.3%
	Non-diagnostic sample	1	7.7%
Distant metastasis		0	0.0%

GIST: Gastrointestinal stromal tumor, EUS: Endoscopic ultrasonography, MRI: Magnetic resonance imaging, PET: Positron emission tomography. Percentages for performed examinations were calculated using the total cohort (n=39), whereas percentages for examination findings were calculated using the number of patients who underwent the corresponding examination

Table 3. Operative characteristics of the patients (n=39)

		Min-max			Median	Mean ± SD/n-%		
Tumor location	Antrum					8		20.5%
	Body					15		38.5%
	Fundus					16		41.0%
Relation to gastric curvature	Greater curvature					13		33.3%
	Lesser curvature					20		51.3%
	Not reported					6		15.4%
Surgical approach	Laparoscopic					30		76.9%
	Open					9		23.1%
Type of gastric resection	Wedge resection					34		87.2%
	Subtotal gastrectomy					3		7.7%
	Total gastrectomy					2		5.1%
Operative time (minutes)		45.0	-	515.0	94.0	114.5	±	85.1

SD: Standard deviation

Table 4. Histopathological characteristics of the patients (n=39)								
		Min-max			Median	Mean ± SD/n-%		
Cell type	Spindle					24		61.5%
	Mixed					8		20.5%
	Epithelioid					5		12.8%
	Not reported					2		5.1%
Cellularity	Low					1		2.6%
	Moderate					3		7.7%
	Marked					25		64.1%
	Not reported					10		25.6%
Cellular atypia	Mild					23		59.0%
	Moderate					3		7.7%
	Marked					2		5.1%
	Not reported					11		28.2%
Tumor size (cm)		0.8	-	32.0	4.5	6.3	±	5.5
Risk group	Very low					17		43.6%
	Low					11		28.2%
	Intermediate					4		10.3%
	High					2		5.1%
	Not reported					5		12.8%
Mitotic count (/5 mm^2)		1.0	-	14.0	2.0	3.0	±	2.8
Necrosis	Absent					23		59.0%
	Present					14		35.9%
	Not reported					2		5.1%
Lymphovascular invasion	Absent					5		12.8%
	Not reported					34		87.2%
Perineural invasion	Absent					3		7.7%
	Not reported					36		92.3%
Resection margin	Negative (R0)					38		97.4%
	Not reported					1		2.6%
Specimen integrity	Complete					24		61.5%
	Incomplete					4		10.3%
	Not reported					11		28.2%

SD: Standard deviation

Patients with Incomplete Specimen Integrity

All four patients with incomplete specimen integrity underwent laparoscopic wedge resection. The intraoperatively estimated tumor diameters ranged from 3.0 to 4.5 cm; these operative estimates were not used for pathological risk classification. Risk categories were based on the pathological tumor diameter and mitotic count reported in the final pathology reports. Two tumors were classified as very low risk and one as intermediate risk; the risk category was not reported for one patient. Resection margins were negative in three patients and unavailable in one. None received adjuvant imatinib.

Recurrence could be assessed from the follow-up records in three of the four patients, and no recurrence was identified. In the remaining patient, vital status at the last follow-up was known; however, the available records were insufficient to evaluate recurrence, and the patient was therefore not considered recurrence-free. All four patients were alive at their last follow-up, with follow-up durations ranging from 8 to 46 months. Although specimen integrity was reported as incomplete in these four patients, tumor rupture could not be confirmed from the available operative notes and detailed pathology records. Therefore, these cases were not classified as definitive tumor rupture.

Immunohistochemical Findings

CD117 expression was positive in 37 of the 38 evaluable patients (97.4%). DOG1 was positive in all 38 evaluable patients (100%), and CD34 was positive in all 36 evaluable patients (100%). SMA staining was positive in 16 of 35 evaluable patients (45.7%), whereas desmin staining was positive in three of 25 evaluable patients (12.0%). S100 was negative in all 33 patients in whom it was evaluated. CD31 was evaluated in only one patient and was negative.

The Ki-67 proliferation index was available for 37 patients. The mean Ki-67 index was 5.4±4.6%, and the median was 5% (range, 1-20%). Among the 37 evaluable patients, the Ki-67 index was <10% in 30 patients (81.1%) and ≥10% in seven (18.9%). Ki-67 results were unavailable for two patients. No patient had an available molecular mutation analysis result for KIT or PDGFRA. The immunohistochemical findings are presented in Table 5.

Postoperative and Medium-term Follow-up Outcomes

The mean postoperative hospital stay was 5.1±3.0 days, with a median of 4 days (range, 2-20 days). No postoperative complication was documented in 33 patients (84.6%), whereas six (15.4%) had a Clavien-Dindo grade II complication. Three patients (7.7%) received adjuvant imatinib.

Ninety-day mortality occurred in one patient (2.6%). Two patients (5.1%) died during the overall follow-up period, while 37 (94.9%) were alive at their last follow-up. The deaths occurred at 2 and 60 months after surgery. One patient (2.6%) died within 90 days after surgery because of progression of concomitant ovarian cancer. This death was unrelated to the gastric GIST resection or a postoperative surgical complication and was therefore not classified as Clavien-Dindo grade V.

Recurrence status was evaluable in 38 patients, none of whom had documented local, peritoneal, or distant GIST recurrence. One patient with insufficient recurrence information was excluded from the denominator. Because no recurrence event occurred, disease-free and recurrence-free survival estimates were not generated.

The mean observed follow-up duration was 41.8±29.8 months, with a median of 33 months (range, 2-120 months). The median follow-up duration estimated using the reverse Kaplan-Meier method was 42 months.

Kaplan-Meier overall survival was 97.4% (95% confidence interval, 83.2-99.6%) at 12, 24, 36, and 48 months and 86.6% (95% confidence interval, 47.5-97.3%) at 60 months; median overall survival was not reached. The numbers at risk at baseline and at 12, 24, 36, 48, 60, 84, and 120 months were 39, 33, 27, 19, 16, 9, 4, and 1, respectively. Because only nine patients remained at risk at 60 months and the confidence interval was wide, this estimate should be interpreted cautiously. Follow-up outcomes are presented in Table 6 and the survival curve in Figure 2.

Discussion

The principal findings of this series were the achievement of R0 resection in all patients with evaluable margin status, the predominance of very low- and low-risk tumors within the cohort, and the absence of documented recurrence among evaluable patients over a median follow-up of 42 months.

The clinical presentation of gastric GISTs varies according to tumor size, anatomical location and relationship with the gastric mucosa. Abdominal pain was the most common presenting symptom in our series, whereas approximately one-fifth of the patients were diagnosed incidentally in the

Table 5. Immunohistochemical findings of the patients (n=39)

Marker	Evaluated n	Positive n (%)	Negative n (%)	Not reported n
CD117	38	37 (97.4)	1 (2.6)	1
DOG1	38	38 (100.0)	0 (0.0)	1
SMA	35	16 (45.7)	19 (54.3)	4
Desmin	25	3 (12.0)	22 (88.0)	14
CD34	36	36 (100.0)	0 (0.0)	3
S100	33	0 (0.0)	33 (100.0)	6
CD31	1	0 (0.0)	1 (100.0)	38
Marker	**Evaluated n**	**Mean ± SD**	**Median**	**Min-max**
Ki-67 (%)	37	5.4±4.6	5.0	1.0-20.0

CD: Cluster of differentiation, DOG1: Discovered on GIST-1, SMA: Smooth muscle actin, SD: Standard deviation, Ki-67: Ki-67 labeling index. Percentages for positive and negative staining were calculated using the number of patients with an available result for the corresponding marker. Molecular testing for KIT and PDGFRA mutations was not performed in any patient

absence of symptoms. Most tumors were located in the fundus or gastric body, and spindle-cell morphology was the predominant histological pattern. The predominance of low mitotic activity and the classification of 82.4% of patients with an available risk category as very low or low risk may partly account for the absence of recurrence during follow-up (1-3,5,10).

Table 6. Postoperative and medium-term follow-up outcomes of the patients (n=39)

		Min-max	Median	Mean ± SD/n-%		
Length of hospital stay (days)		2.0-20.0	4.0	5.1	±	3.0
Postoperative complication status	No postoperative complication			33		84.6%
	Grade II			6		15.4%
Adjuvant imatinib therapy				3		7.7%
Tumor recurrence				0		0.0%
90-day mortality				1		2.6%
Overall mortality				2		5.1%
Observed follow-up (months)		2.0-120.0	33.0	41.8	±	29.8
Median follow-up by reverse Kaplan-Meier, months			42.0			
24-month overall survival			97.4% (95% CI, 83.2-99.6)			
48-month overall survival			97.4% (95% CI, 83.2-99.6)			
60-month overall survival			86.6% (95% CI, 47.5-97.3)			

Data are presented as mean ± standard deviation (SD), median (minimum-maximum), n (%), n/evaluable n (%), or Kaplan-Meier estimate (95% confidence interval), as appropriate. Recurrence status was available for 38 patients; one patient with missing recurrence information was excluded from the denominator. Because no documented recurrence occurred, disease-free survival and recurrence-free survival were not estimated. Median follow-up was calculated using the reverse Kaplan-Meier method. Only nine patients remained at risk at 60 months; therefore, the 60-month overall survival estimate should be interpreted cautiously.
CI: Confidence interval

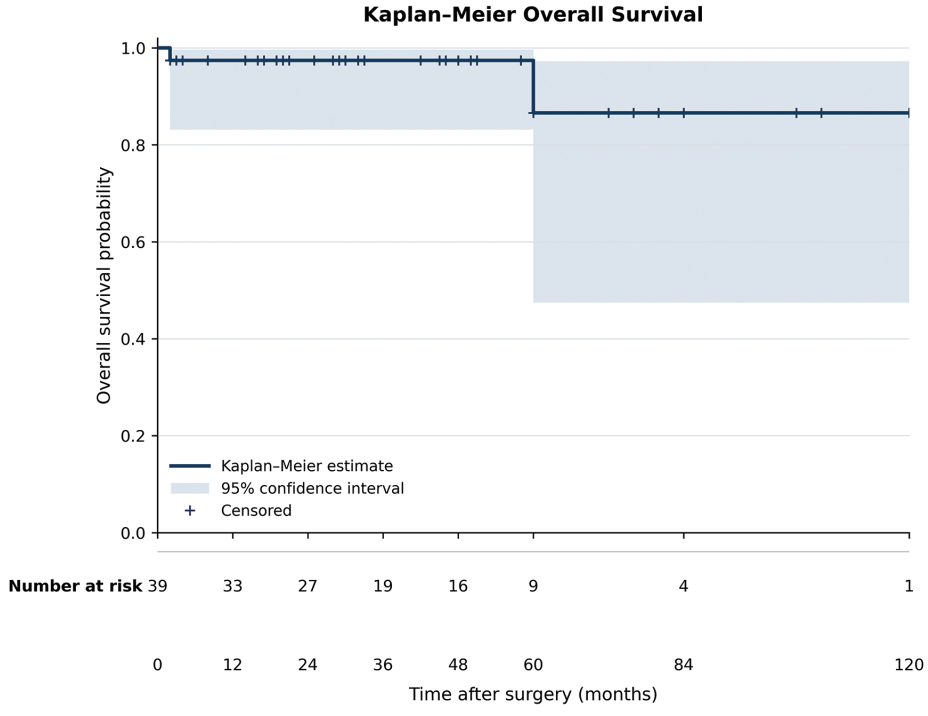

Figure 2. Kaplan-Meier estimate of overall survival after surgical resection for gastric gastrointestinal stromal tumor. The shaded area represents the 95% confidence interval, and vertical marks indicate censored observations. Numbers at risk are presented below the curve

The primary objective of surgery for localized and resectable gastric GIST is to achieve macroscopically complete resection with a microscopically negative margin while preserving the tumor pseudocapsule. The anatomical characteristics of the stomach allow organ-preserving wedge resection to be performed instead of more extensive gastric resection in appropriately located tumors (1-3). In our series, wedge resection was performed in 87.2% of patients, and R0 resection was achieved in all patients with evaluable margin status. These findings indicate that the primary surgical objective was achieved in all patients with evaluable margin status.

A laparoscopic approach was used in 76.9% of the patients. Current guidelines consider minimally invasive surgery appropriate for selected gastric GISTs when tumor integrity and oncological principles can be maintained (1-3). Recent multicenter and national database studies have associated minimally invasive surgery with shorter hospital stay and favorable perioperative outcomes, while suggesting comparable oncological outcomes in appropriately selected patients (7,8,13). However, these studies are largely retrospective, and patients undergoing open surgery generally have larger or technically more complex tumors, creating selection bias. Because only nine patients in our cohort underwent open surgery and the approach was not randomized, no assessment of superiority was made. The high laparoscopy rate should therefore be interpreted as institutional patient selection and practice rather than evidence of the superiority of one approach.

The median hospital stay of four days was generally consistent with contemporary gastric GIST surgery series (8,13). According to the available records, no postoperative complication was documented in 33 patients, whereas six had a Clavien-Dindo grade II complication. The patient who died within 90 days died from progression of concomitant ovarian cancer rather than a surgical complication and was therefore not classified as Clavien-Dindo grade V. However, limited retrospective information regarding the type, treatment, and duration of the grade II complications precluded a comprehensive assessment of perioperative outcomes.

Tumor rupture is an independent and clinically important risk factor for recurrence and may directly influence risk classification and decisions regarding adjuvant therapy (1-3,11). Specimen integrity was reported as incomplete in four patients; however, tumor rupture could not be confirmed from the available operative and pathological records. Accordingly, these cases were not classified as definitive tumor rupture. None of these four patients received adjuvant imatinib, and no documented recurrence was identified in the three patients with evaluable recurrence data. Nevertheless, the small number of patients, missing recurrence information for one patient, and limited follow-up durations of 8-46 months preclude any prognostic interpretation. This observation emphasizes the importance of documenting tumor integrity and any potential rupture event using standardized definitions at the time of surgery (11).

Among evaluable patients, CD117 positivity was 97.4%, whereas DOG1 and CD34 were positive in 100%. These results are consistent with the established immunophenotype of gastric GISTs and support the histopathological diagnoses (1-3,5), but they are expected diagnostic findings rather than an original contribution. The limited assessment of lymphovascular and perineural invasion and missing information on risk classification and specimen integrity restricted the pathological assessment. Current guidelines recommend standardized reporting of molecular characteristics together with tumor size, anatomical location, mitotic activity, margin status, and tumor rupture (1-3,14).

Study Limitations

The absence of KIT and PDGFRA mutation analysis was a major limitation. Molecular classification is a central component of contemporary GIST management because it informs imatinib sensitivity, identifies subgroups with primary resistance, and guides tyrosine kinase inhibitor selection (1-4,6,14,15). Certain PDGFRA mutations are associated with primary imatinib resistance, whereas the type of KIT mutation may affect the expected treatment benefit (1-3,15). Although three patients received adjuvant imatinib, its molecular appropriateness could not be assessed. Long-term trial data support three years of adjuvant imatinib in patients with high-risk, imatinib-sensitive GIST (15,16); however, these findings cannot be directly applied to our patients because mutation status, treatment duration, and adherence were incompletely documented.

Variability in follow-up duration was addressed using the reverse Kaplan-Meier method. Because no documented recurrence occurred among evaluable patients, a recurrence-free survival curve was not presented; only deaths from any cause were treated as events in the overall survival analysis.

The predominance of very low- and low-risk tumors within the cohort may partly explain the absence of recurrence during follow-up. Nevertheless, late recurrence has been reported even in patients with low-risk GIST. In a large series from the Italian Sarcoma Group, recurrence occurred in 5.7% of patients with low-risk GIST (17), whereas Berndsen et al. (18) reported highly favorable long-term disease-specific outcomes after complete resection of non-high-risk GISTs. Thus, although our findings support favorable medium-term oncological control, they do not exclude the possibility of late recurrence.

The five-year overall survival estimate should be interpreted cautiously because only nine patients remained at risk at 60 months and the corresponding confidence interval was wide. Moreover, overall survival included deaths from any cause; therefore, the observed deaths could not be attributed directly to GIST. For these reasons, the findings should be regarded as medium-term real-world outcomes rather than evidence of long-term survival.

The principal contribution of this study is its description of organ-preserving surgical practice and medium-term outcomes in a consecutive cohort of patients with histopathologically confirmed gastric GIST, supported by an appropriate follow-up analysis. The retrospective single-center design, limited sample, missing pathological data, and absence of molecular testing restrict generalizability. Conversely, the inclusion of only confirmed gastric GISTs, transparent use of evaluable denominators, and presentation of survival estimates with the corresponding numbers at risk are methodological strengths.

Conclusion

This single-center series demonstrates that organ-preserving wedge resection was frequently feasible in selected patients with localized gastric GIST and that R0 resection was achieved in all patients with evaluable margin status. Although no documented recurrence was identified during medium-term follow-up, the predominance of low-risk tumors, limited sample size, and declining numbers of patients at risk at later time points preclude firm conclusions regarding long-term oncological effectiveness. These findings support individualizing the surgical approach and extent of resection according to tumor anatomy and the feasibility of preserving tumor integrity, together with standardized documentation of pathological and molecular data.

Ethics

Ethics Committee Approval: The study was approved by the Local Clinical Research Ethics Committee of University of Health Sciences Turkey, İstanbul Bakırköy Dr. Sadi Konuk Training and Research Hospital before its initiation (approval date: June 19, 2026; decision no: 2026-14-26).

Informed Consent: Written informed consent for the use of clinical data for research purposes was obtained from all patients before their inclusion in the study.

Acknowledgments

Artificial intelligence (ChatGPT, OpenAI) was used solely for language editing, grammar correction, and improvement of the academic writing style. No artificial intelligence-generated content was used in the study design, data collection, statistical analysis, interpretation of the results, or scientific decision-making. All scientific content was reviewed, verified, and approved by the authors.

Footnotes

Authorship Contributions

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

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

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

References

1. Serrano C, Martín-Broto J, Asencio-Pascual JM, López-Guerrero JA, Rubió-Casadevall J, Bagué S, et al. 2023 GEIS guidelines for gastrointestinal stromal tumors. Ther Adv Med Oncol. 2023;15:17588359231192388.
2. Casali PG, Blay JY, Abecassis N, Bajpai J, Bauer S, Biagini R, et al. Gastrointestinal stromal tumours: ESMO-EURACAN-GENTURIS Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2022;33(1):20-33.
3. Hirota S, Tateishi U, Nakamoto Y, Yamamoto H, Sakurai S, Kikuchi H, et al. English version of Japanese Clinical Practice Guidelines 2022 for gastrointestinal stromal tumor (GIST) issued by the Japan Society of Clinical Oncology. Int J Clin Oncol. 2024;29(6):647-680.
4. von Mehren M, Kane JM, Riedel RF, Sicklick JK, Pollack SM, Agulnik M, et al. NCCN Guidelines® insights: gastrointestinal stromal tumors, version 2.2022. J Natl Compr Canc Netw. 2022;20(11):1204-1214.

5. Blay JY, Kang YK, Nishida T, von Mehren M. Gastrointestinal stromal tumours. Nat Rev Dis Primers. 2021;7(1):22.

6. Serrano C, Álvarez R, Carrasco JA, Marquina G, Martínez-García J, Martínez-Marín V, et al. SEOM-GEIS clinical guideline for gastrointestinal stromal tumors (2022). Clin Transl Oncol. 2023;25(9):2707-2717.

7. Park SH, Lee HJ, Kim MC, Yook JH, Sohn TS, Hyung WJ, et al. Early experience of laparoscopic resection and comparison with open surgery for gastric gastrointestinal stromal tumor: a multicenter retrospective study. Sci Rep. 2022;12(1):2290.

8. Crocker AB, Vega EA, Kutlu OC, Salehi O, Mellado S, Li M, et al. Is minimally invasive surgery for large gastric GIST actually safe? A comparative analysis of short- and long-term outcomes. Surg Endosc. 2022;36(9):6975-6983.

9. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. Ann Intern Med. 2007;147(8):573-577.

10. Joensuu H. Risk stratification of patients diagnosed with gastrointestinal stromal tumor. Hum Pathol. 2008;39(10):1411-1419.

11. Nishida T, Hølmebakk T, Raut CP, Rutkowski P. Defining tumor rupture in gastrointestinal stromal tumor. Ann Surg Oncol. 2019;26(6):1669-1675.

12. Schemper M, Smith TL. A note on quantifying follow-up in studies of failure time. Control Clin Trials. 1996;17(4):343-346.

13. Freeman HD, Mudgway R, Tran Z, Kim R, Lum SS, Namm JP, et al. Oncologic outcomes and survival of modern surgical approaches for gastric gastrointestinal stromal tumor (GIST). Surg Endosc. 2024;38(11):6854-6864.

14. Judson I, Jones RL, Wong NACS, Dileo P, Bulusu R, Smith M, et al. Gastrointestinal stromal tumour (GIST): British sarcoma group clinical practice guidelines. Br J Cancer. 2025;132(1):1-10.

15. Joensuu H, Wardelmann E, Eriksson M, Reichardt A, Sundby Hall K, Schütte J, et al. KIT and PDGFRA mutations and survival of gastrointestinal stromal tumor patients treated with adjuvant imatinib in a randomized trial. Clin Cancer Res. 2023;29(17):3313-3319.

16. Joensuu H, Eriksson M, Sundby Hall K, Reichardt A, Hermes B, Schütte J, et al. Survival outcomes associated with 3 years vs 1 year of adjuvant imatinib for patients with high-risk gastrointestinal stromal tumors: an analysis of a randomized clinical trial after 10-year follow-up. JAMA Oncol. 2020;6(8):1241-1246.

17. D'Ambrosio L, Fumagalli E, De Pas TM, Nannini M, Bertuzzi A, Carpano S, et al. Guideline-based follow-up outcomes in patients with gastrointestinal stromal tumor with low risk of recurrence: a report from the Italian Sarcoma Group. JAMA Netw Open. 2023;6(11):e2341522.

18. Berndsen M, Renberg S, Hølmebakk T, Hancke E, Puls F, Karlsson F, et al. Long-term outcome after surgical resection of non-high-risk gastrointestinal stromal tumours without adjuvant therapy. Br J Surg. 2023;110(12):1857-1862.

CASE REPORT

Bagcilar Med Bull 2026;11(3):427-430
DOI: 10.4274/BMB.galenos.2026.51422

Occult Bronchial Carcinoid Tumor: A Rare Case of Lung Cancer

Okült Bronşiyal Karsinoid Tümör: Nadir Bir Akciğer Kanseri Olgusu

Hülya Abalı, Ayşenur İlkyaz, Çiğdem Sabancı, Sida Gösterici, Fatma Tokgöz Akyıl

University of Health Sciences Turkey, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital, Department of Pulmonology, İstanbul, Turkey

Abstract

Bronchial carcinoid tumors are uncommon malignant neuroendocrine tumors that account for 2% of lung cancers. The main diagnostic tools are contrast-enhanced thoracic computed tomography (CT) and bronchoscopy. Early diagnosis is crucial because surgical excision is the primary treatment. The authors report a 32-year-old male patient with the smallest typical bronchial carcinoid tumor reported in the literature. The patient presented with chest pain for one month and recurrent pneumonia. Thoracic CT revealed an unclear 1 cm perihilar mass. Positron emission tomography CT showed a showing no ^{18}F-FDG uptake 1.4 cm nodule in the superior segment of the right lower lobe. The tumor was visualized via fiberoptic bronchoscopy, and initial biopsies were taken. Definitive diagnosis and curative treatment were achieved via segmentectomy. Bronchoscopy should be performed on a 32-year-old man patients without comorbidities who experience recurrent pneumonia or have small hilar or perihilar masses on imaging.

Keywords: Bronchial carcinoid, bronchoscopy, contrast-enhanced CT, FDG PET, typical carcinoid

Öz

Bronşiyal karsinoid tümörler, nadir görülen malign nöroendokrin tümörlerdir ve akciğer kanserlerinin %2'sini oluşturur. Başlıca tanı araçları, kontrastlı toraks bilgisayarlı tomografi (BT) ve bronkoskopidir. Erken tanı önemlidir, çünkü cerrahi eksizyon temel tedavi yöntemidir. Yazarlar, literatürdeki en küçük tipik bronşiyal karsinoid tümörü olan 32 yaşında bir erkek hastayı sunmaktadır. Hasta bir aydır göğüs ağrısı ve reküren pnömoni şikayetleriyle başvurdu. Toraks BT'sinde belirsiz 1 cm çapında perihiler kitle görüldü. Pozitron emisyon tomografisi BT'sinde ise sağ alt lob superior segmentinde ^{18}F-FDG tutulumu göstermeyen 1,4 cm büyüklüğünde bir nodül saptandı. Fiberoptik bronkoskopi ile tümör görüntülendi ve ilk biyopsiler alındı. Kesin tanı ve tedavi segmentektomi ile sağlandı. Orta yaş grubundaki, komorbiditeleri olmayan, reküren pnömoni gelişen veya görüntülemesinde küçük hiler ve perihiler kitle görülen hastalara bronkoskopi yapılmalıdır.

Anahtar kelimeler: Bronkoskopi, bronşiyal karsinoid, FDG PET, kontrastlı BT, tipik karsinoid

Introduction

Bronchial carcinoid tumors (BCTs) are rare, solitary, slow-growing malignancies arising from neuroendocrine cells of the bronchial epithelium (1). Segmental and subsegmental BCTs constitute 80-90% of all cases (2). Well-differentiated BCTs are categorized into low-grade typical carcinoids (TCs), which have a favorable prognosis, and intermediate-grade atypical carcinoids (ACs), which carry a poorer prognosis (3). TCs are more common than ACs, with a ratio of 8-10:1 (4).

Most patients present with symptoms such as cough, dyspnea, or recurrent pneumonia; however, 25-39% remain asymptomatic due to the indolent nature of these tumors (5,6).

Address for Correspondence: Hülya Abalı, University of Health Sciences Turkey, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital, Department of Pulmonology, İstanbul, Turkey
E-mail: hulayab@gmail.com **ORCID:** orcid.org/0000-0003-4041-7479
Received: 23.03.2025 **Accepted:** 01.02.2026 **Epub:** 03.02.2026 **Publication Date:** 29.09.2026

Cite this article as: Abalı H, İlkyaz A, Sabancı Ç, Gösterici S, Tokgöz Akyıl F. Occult bronchial carcinoid tumor: a rare case of lung cancer. Bagcilar Med Bull. 2026;11(3):427-430

Contrast-enhanced computed tomography (CT) is the primary imaging modality for detecting BCTs. Common CT findings include perihilar or hilar masses, endobronchial nodules, and signs of bronchial obstruction—such as atelectasis, post-obstructive pneumonitis, or air trapping (7,8). Perihilar and hilar masses may also be associated with hilar or mediastinal lymphadenopathy secondary to metastasis or recurrent infection (7).

Definitive diagnosis relies on histopathological and immunohistochemical evaluation of tissue samples. Bronchoscopy remains an essential tool for confirming diagnosis, assessing bronchial wall invasion, and enabling therapeutic removal of endobronchial lesions, which may resolve associated atelectasis (9).

Here, we present what to our knowledge is the smallest TC reported in the literature—unrecognized as a tumor initially on non-contrast CT yet clearly visualized by fiber-optic bronchoscopy (FOB).

Case Report

A 32-year-old man was referred to our pulmonology clinic for evaluation of recurrent pneumonia over the preceding year. He reported one month of chest pain. His medical history was unremarkable, and he smoked three cigarettes daily for 10 years. Family history included an unspecified malignancy in a grandparent. Physical examination and laboratory tests were normal.

Non-contrast CT performed at an outside facility revealed a 1-cm right perihilar lesion, initially suspected to represent a lymph node (Figure 1), prompting a positron emission tomography-CT (PET-CT). PET-CT demonstrated a small, slightly hyperdense 1.4-cm nodular lesion in the superior segment of the right lower lobe (RLL) without ^{18}F-fluorodeoxyglucose (^{18}F-FDG) uptake (Figure 2). Contrast-enhanced brain magnetic resonance imaging showed no metastatic disease.

FOB revealed near-complete obliteration of the superior segmental bronchus of the RLL by a smooth, hypervascular, cherry-red polypoid mass (Figure 3). Multiple punch biopsies were obtained, and histopathology indicated a carcinoid tumor. Due to an inadequate sample, the carcinoid subtype was reported as suspected typical carcinoid. Surgery was planned to confirm the diagnosis and for treatment. The patient subsequently underwent segmentectomy with bronchoplasty after multidisciplinary surgical review. Pathology confirmed a 1.4×1.1×0.9 cm TC. The sampled right hilar and infrahilar lymph nodes showed no involvement. No recurrence was detected during three years of postoperative follow-up, and he remains under annual surveillance.

Discussion

BCTs are more common in females and white patients, affecting individuals across all age groups. Mean age at diagnosis is approximately 55 years for ACs and 45 years for TCs (10). Although BCTs are not classically linked to carcinogen exposure, smoking is associated with their occurrence (11). Our patient's demographic profile—a

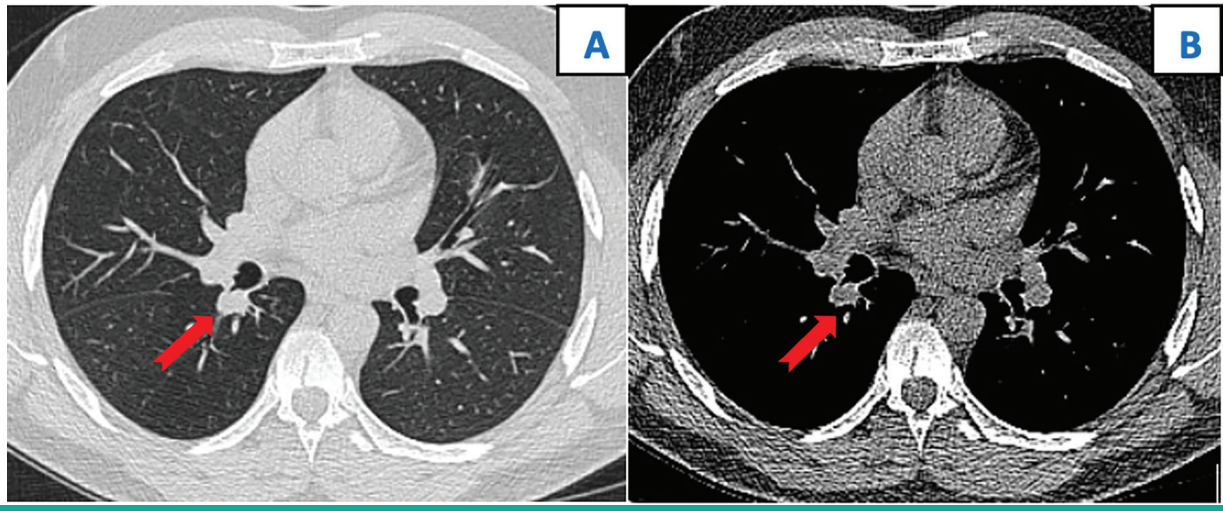

Figure 1. A 1 cm lesion in the right lung perihilar area on a non-contrast CT scan taken at an external center (b). A: Parenchymal image, B: Mediastinal image

CT: Computed tomography

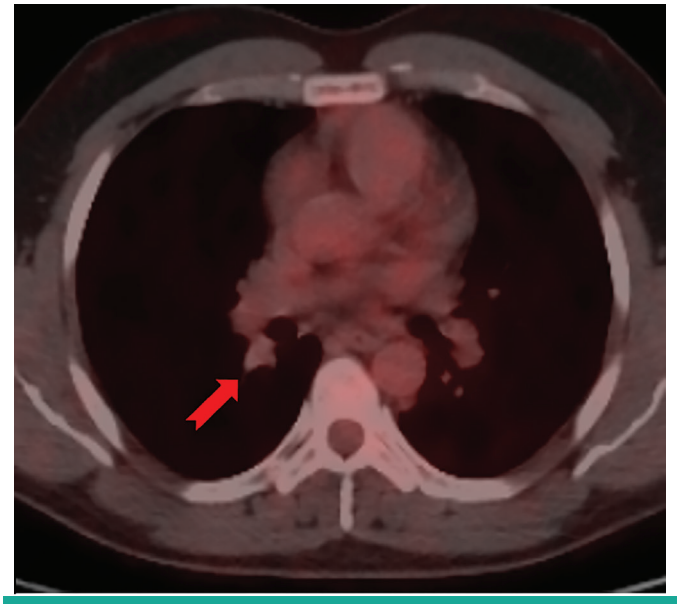

Figure 2. A 1.4 cm nodule in the right lung lower lobe superior segment with no 18F-fluorodeoxyglucose uptake on PET-CT (indicated by a red arrow)

PET-CT: Positron emission tomography-computed tomography

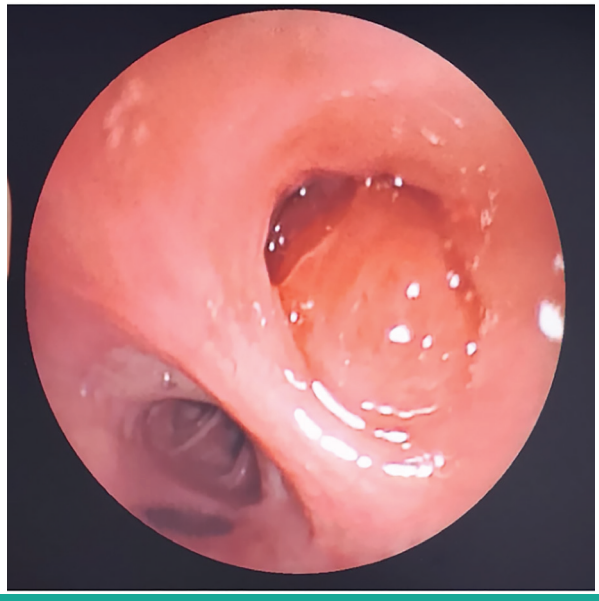

Figure 3. Bronchoscopy image of a smooth-surfaced, cherry-red, polypoid, typical carcinoid tumor obliterating the entrance of the right lung lower lobe superior segment bronchus

32-year-old man, white, and a smoker—aligns with known epidemiologic patterns.

Typical symptoms include chest pain, recurrent pneumonia, hemoptysis, and wheezing due to bronchial obstruction (7). Carcinoid syndrome is rare and typically associated with metastatic disease. The mediastinal lymph nodes are the most common metastatic site (12). Consistent with prior studies, the patient with localized TC presented with chest pain and recurrent pneumonia without features of carcinoid syndrome. The diagnosis is often delayed, even when symptoms are apparent. For example, patients with recurrent pneumonia may undergo years of evaluations before receiving a definitive diagnosis (13). Similarly, our patient sought care at multiple facilities for recurrent pneumonia and was treated with several courses of antibiotics over the span of a year before the TC was ultimately identified at our center.

TCs commonly appear as well-circumscribed ovoid or round opacities on CT, with 75-77% located centrally in the main, lobar, or segmental bronchi (14). TCs commonly arise in the right lung and middle lobe (13). The tumor in this case originated in the segmental bronchus of the RLL, consistent with these findings. Most BCTs measure 2-5 cm (7), whereas the tumor in our patient was the smallest reported, with a diameter of only 1.4 cm. Although contrast-enhanced CT can detect small carcinoids, the initial non-contrast CT contributed to diagnostic difficulty.

^{18}F-FDG PET-CT is a widely used imaging technique to assess thoracic cancers. PET-CT generally has limited value for well-differentiated BCTs, with sensitivity increasing with higher grade, from TCs to ACs (15). The absence of ^{18}F-fluorodeoxyglucose uptake in this case aligns with expectations for TC.

Bronchoscopy remains crucial for evaluating BCTs, enabling direct visualization and biopsy. The tumor's appearance—hypervascular, fragile, and polypoid—is characteristic of TC (9). Due to inadequate sampling and small size, more than 30% of BCTs require surgical resection for definitive diagnosis (16), as occurred in this patient.

Early detection is vital because surgical resection provides excellent outcomes in localized BCTs (17). Although TCs are less aggressive than ACs, they metastasize to regional lymph nodes in 4-20% of cases (18). Prognosis correlates more strongly with nodal status than with histologic subtype; 5-year survival reaches 100% in N0 patients for both TC and AC, while it decreases to 90% and 78.8% in N1 patients for TC and AC, respectively (19). Tumor size <3 cm also correlates with better outcomes (20). The patient with a small N0 TC underwent a segmentectomy and experienced an uneventful postoperative course with no recurrence at three years.

Conclusion

Clinicians should be aware of BCTs in middle-aged smokers without comorbidities presenting with recurrent pneumonia. Early diagnosis of BCTs is crucial, as surgical excision is curative for localized carcinoids. Contrast-enhanced thoracic CT is the mainstay of imaging for detecting carcinoids; however, small TCs may be radiologically occult. We recommend that physicians carefully examine the hilar and perihilar areas on CT scans and perform FOB to identify potential endobronchial tumors.

Ethics

Informed Consent: Written informed consent was obtained from the patient for the publication of the article and the images related to the report.

Footnotes

The case report was presented as an E-poster at the 27th Annual Congress of the Turkish Thoracic Organization (29 April-3 May 2024, Girne).

Authorship Contributions

Surgical and Medical Practices: H.A., Concept: H.A., A.İ., Ç.S., S.G., F.T.A., Design: H.A., Data Collection or Processing: H.A., A.İ., Ç.S., S.G., F.T.A., Analysis or Interpretation: H.A., A.İ., Ç.S., S.G., F.T.A., Literature Search: H.A., Writing: H.A.

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

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

References

1. Borczuk AC. Pulmonary neuroendocrine tumors. Surg Pathol Clin. 2020;13(1):35-55.
2. Hage R, de la Rivière AB, Seldenrijk CA, van den Bosch JMM. Update in pulmonary carcinoid tumors: a review article. Ann Surg Oncol. 2003;10(6):697-704.
3. Pericleous M, Karpathakis A, Toumpanakis C, Lumgair H, Reiner J, Marelli L, et al. Well-differentiated bronchial neuroendocrine tumors: clinical management and outcomes in 105 patients. Clin Respir J. 2018;12(3):904-914.
4. Caplin ME, Baudin E, Ferolla P, Filosso P, Garcia-Yuste M, Lim E, et al. ENETS consensus conference participants. Pulmonary neuroendocrine (carcinoid) tumors: European Neuroendocrine Tumor Society expert consensus and recommendations for best practice for typical and atypical pulmonary carcinoids. Ann Oncol. 2015;26(8):1604-1620.
5. van der Heijden EHFM. Bronchial carcinoid? Interventional pulmonologist first! Respiration. 2018;95(4):217-219.
6. Chong S, Lee KS, Chung MJ, Han J, Kwon OJ, Kim TS. Neuroendocrine tumors of the lung: clinical, pathologic, and imaging findings. Radiographics. 2006;26(1):41-57; discussion 57-58.
7. Jeung MY, Gasser B, Gangi A, Charneau D, Ducroq X, Kessler R, et al. Bronchial carcinoid tumors of the thorax: spectrum of radiologic findings. Radiographics. 2002;22(2):351-365.
8. Strange CD, Strange TA, Erasmus LT, Patel S, Ahuja J, Shroff GS, et al. Imaging in lung cancer staging. Clin Chest Med. 2024;45(2):295-305.
9. Papaporfyriou A, Domayer J, Meilinger M, Firlinger I, Funk GC, Setinek U, et al. Bronchoscopic diagnosis and treatment of endobronchial carcinoid: case report and review of the literature. Eur Respir Rev. 2021;30(159):200115.
10. Brisset C, Roumy M, Lacour B, Hescot S, Bras ML, Dijoud F, et al. Bronchial carcinoid tumors in children and adolescents. Pediatr Blood Cancer. 2025;72(8):e31822.
11. Mindaye ET, Kassahun M, Tigiye G. Bronchial carcinoid tumor in the era of covid-19 pandemic: a case report. Int J Surg Case Rep. 2021;80:105703.
12. Kapoor R, Mandelia A, Farzana N, Nigam N, Dabadghao P, Sharma SP, et al. Bronchial carcinoid tumor in an adolescent female: diagnosis and management by a multi-disciplinary team. J Indian Assoc Pediatr Surg. 2022;27(4):500-502.
13. Fink G, Krelbaum T, Yellin A, Bendayan D, Saute M, Glazer M, et al. Pulmonary carcinoid: presentation, diagnosis, and outcome in 142 cases in Israel and review of 640 cases from the literature. Chest. 2001;119(6):1647-1651.
14. Kaifi JT, Kayser G, Ruf J, Passlick B. The diagnosis and treatment of bronchopulmonary carcinoid. Dtsch Arztebl Int. 2015;112(27-28):479-485.
15. Zhao J, Wang H. Differential diagnostic value of 18F-FDG PET/CT in pulmonary carcinoids versus hamartomas. Acad Radiol. 2022;29(Suppl 2):S41-S46.
16. Detterbeck FC. Management of carcinoid tumors. Ann Thorac Surg. 2010;89(3):998-1005.
17. Reuling EMBP, Dickhoff C, Plaisier PW, Bonjer HJ, Daniels JMA. Endobronchial and surgical treatment of pulmonary carcinoid tumors: a systematic literature review. Lung Cancer. 2019;134:85-95.
18. Morandi U, Casali C, Rossi G. Bronchial typical carcinoid tumors. Semin Thorac Cardiovasc Surg. 2006;18(3):191-198.
19. Cardillo G, Sera F, Di Martino M, Graziano P, Giunti R, Carbone L, et al. Bronchial carcinoid tumors: nodal status and long-term survival after resection. Ann Thorac Surg. 2004;77(5):1781-1785.
20. Yang Z, Wang Z, Duan Y, Xu S. Clinicopathological characteristics and prognosis of resected cases of carcinoid tumors of the lung. Thorac Cancer. 2016;7(6):633-638.