E-ISSN: 2147-2688

The Medical Bulletin of Haseki

2026
Volume 64
Issue
November 5

www.hasekidergisi.com

The Medical Bulletin of Haseki

Editorial Board

Owner - On behalf of University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital,

Prof. Dr. Mehmet Mesut Sonmez (Chief Physician of Haseki Training and Research Hospital)
University of Health Sciences Türkiye, Hamidiye Faculty of Medicine, Clinic of Orthopedics and Traumatology, Istanbul, Türkiye
E-mail: mdmesutsonmez@yahoo.com

Editors-in-Chief

Assoc. Prof. Akif Erbin (Surgical Medical Sciences)
University of Health Sciences Türkiye, Hamidiye Faculty of Medicine, Basaksehir Cam and Sakura City Hospital, Clinic of Urology, Istanbul, Türkiye
E-mail: akiferbin@hotmail.com

Assoc. Prof. Egemen Cebeci (Internal Medical Sciences)
University of Health Sciences Türkiye, Hamidiye Faculty of Medicine, Istanbul Haseki Training and Research Hospital, Clinic of Nephrology, Istanbul, Türkiye
E-mail: dregemencebeci@hotmail.com

Managing Editors

Prof. Dr. Mahmut Edip Gurol
Harvard Medical School, Massachusetts General Hospital, Department of Neurology, Boston, USA
E-mail: edip@mail.harvard.edu

Prof. Dr. Erdem Tuzun
Istanbul University, Aziz Sancar Department of Institute of Experimental Medicine, Department of Neuroscience, Istanbul, Türkiye
E-mail: erdemtuzun@istanbul.edu.tr

Prof. Dr. Elif Cadirci
Ataturk University Faculty of Medicine, Department of Medical Pharmacology, Erzurum, Türkiye
E-mail: ecadirci@atauni.edu.tr

Prof. Dr. Ali Ahiskalioglu
Ataturk University Faculty of Medicine, Department of Anesthesiology and Reanimation, Erzurum, Türkiye
E-mail: aliahiskalioglu@hotmail.com

Statistical Editor

Ahmet Dirican
Istanbul University, Istanbul Faculty of Medicine, Department of Biostatistics and Medical Informatics, Istanbul, Türkiye

Reviewing the articles' conformity to the publishing standards of the Journal, typesetting, reviewing and editing the manuscripts and abstracts in English, creating links to source data, and publishing process are realized by Galenos.

All rights are reserved. Rights to the use and reproduction, including in the electronic media, of all communications, papers, photographs and illustrations appearing in this journal belong to the Medical Bulletin of University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital. Reproduction without prior written permission of part or all of any material is forbidden. The journal complies with the Professional Principles of the Press.

Publisher Contact
Address: Molla Gurani Mah. Kacamak Sk. No: 21/1 34093 Istanbul, Türkiye
Phone: +90 (530) 177 30 97 / +90 (539) 307 32 03
E-mail: info@galenos.com.tr/yayin@galenos.com.tr
Web: www.galenos.com.tr
Publisher Certificate Number: 14521

Online Publishing Date: November 2026
E-ISSN: 2147-2688
International scientific journal published quarterly.

The Medical Bulletin of Haseki

Scientific Advisory Board

Richard J Johnson
Colorado University Anschutz Medical Campus, Department of Renal Diseases and Hypertension, Aurora, Colorado, USA

David Goldsmith
Professor and Emeritus Consultant Nephrologist, Renal Unit, Guy's and St Thomas' Hospital, London, UK

Adrian Covic
Grigore T. Popa University of Medicine and Pharmacy, Department of Internal Medicine (Nephrology), Iasi, Romania

Rumeyza Kazancioglu
Bezmialem Vakif University Faculty of Medicine, Department of Internal Medicine, Division of Nephrology, Istanbul, Türkiye

Alaaddin Yildiz
Istanbul University, Istanbul Faculty of Medicine, Department of Internal Medicine, Division of Nephrology, Istanbul, Türkiye

Suleyman Tevfik Ecder
Istanbul Science University Faculty of Medicine, Department of Internal Medicine, Division of Nephrology, Istanbul, Türkiye

Mehmet Kanbay
Koc University Faculty of Medicine, Department of Internal Medicine, Division of Nephrology, Istanbul, Türkiye

Sule Poturoglu
University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Internal Medicine, Division of Gastroenterology, Istanbul, Türkiye

Evrim Cakir
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of Internal Medicine, Division of Endocrinology, Istanbul, Türkiye

Ozgur Tanriverdi
Mugla Sitki Kocman University Faculty of Medicine, Department of Internal Diseases, Division of Medical Oncology, Mugla, Türkiye

Hayriye Esra Ataoglu
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of General Internal Medicine, Istanbul, Türkiye

Zeynep Karaali
University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of General Internal Medicine, Istanbul, Türkiye

Mustafa Yenigun
Medical Park Hospital, Clinic of General Internal Medicine, Istanbul, Türkiye

Abdulbaki Kumbasar
University of Health Sciences Türkiye, Kanuni Sultan Suleyman Training and Research Hospital, Clinic of General Internal Medicine, Istanbul, Türkiye

Mustafa Yildirim
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of Infectious Diseases and Clinical Microbiology, Istanbul, Türkiye

Aydin Alper
Koc University Faculty of Medicine, Department of General Surgery, Istanbul, Türkiye

Vahit Ozmen
Istanbul University, Istanbul Faculty of Medicine, Department of General Surgery, Istanbul, Türkiye

Gokcen Orhan
University of Health Sciences Türkiye, Dr. Siyami Ersek Thoracic and Cardiovascular Surgery Training and Research Hospital, Clinic of Cardiovascular Surgery, Istanbul, Türkiye

Deniz Goksedef
Istanbul University-Cerrahpasa, Cerrahpasa Faculty of Medicine, Department of Cardiovascular Surgery, Istanbul, Türkiye

Jose L. Peiró
Cincinnati University Faculty of Medicine, Department of Pediatric General and Thoracic Surgery, Cincinnati, USA

Celalettin Ibrahim Kocaturk
University of Health Sciences Türkiye, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital, Clinic of Chest Surgery, Istanbul, Türkiye

Soner Duru
Duzce University Faculty of Medicine, Department of Brain and Nerve Surgery, Duzce, Türkiye

Pakizer Banu Kilicoglu Dane
Bezmialem Vakif University Faculty of Medicine, Department of Obstetrics and Gynecology, Istanbul, Türkiye

Ates Kadioglu
Istanbul University, Istanbul Faculty of Medicine, Department of Urology, Istanbul, Türkiye

The Medical Bulletin of Haseki

Scientific Advisory Board

Ahmet Yaser Muslumanoglu
University of Health Sciences Türkiye, Bagcilar Training and Research Hospital, Clinic of Urology, Istanbul, Türkiye

Murat Binbay
Hasan Kalyoncu University Faculty of Medicine, Department of Urology, Istanbul, Türkiye

Orhan Ozturan
Bezmialem Vakif University Faculty of Medicine, Department of Otorhinolaryngology, Istanbul, Türkiye

Husamettin Yasar
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of Otorhinolaryngology, Istanbul, Türkiye

Fadlullah Aksoy
Bezmialem Vakif University Faculty of Medicine, Department of Otorhinolaryngology, Istanbul, Türkiye

Murat Haluk Ozkul
Istinye University Medical Park Gaziosmanpasa Hospital, Department of Otorhinolaryngology, Istanbul, Türkiye

Fatma Nilufer Alparslan Sansoy
Istanbul University, Istanbul Faculty of Medicine, Department of Ophthalmology, Istanbul, Türkiye

Mahmut Ercan Cetinus
Istinye University Medical Park Gaziosmanpasa Hospital, Department of Otorhinolaryngology, Istanbul, Türkiye

Irfan Ozturk
Florence Nightingale Hospital, Clinic of Orthopedics and Traumatology, Istanbul, Türkiye

Lutfi Telci
Acibadem Hospital, Clinic of Anesthesia and Reanimation, Istanbul, Türkiye

Birgul Bastan Tuzun
University of Health Sciences Türkiye, Hamidiye Faculty of Medicine, Istanbul Haseki Training and Research Hospital, Department of Neurology, Istanbul, Türkiye

Ayse Ozlem Cokar
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of Neurology, Istanbul, Türkiye

Murat Elevli
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of Child Health and Diseases, Istanbul, Türkiye

Saliha Senel
Ankara Yildirim Beyazit University Faculty of Medicine, Department of Child Health and Diseases, Ankara, Türkiye

Mahmut Civilibal
Memorial Bahcelievler Hospital, Clinic of Pediatric Nephrology, Istanbul, Türkiye

Bulent Enis Sekerel
Hacettepe University Faculty of Medicine, Department of Child Health and Diseases, Division of Pediatric Allergy and Asthma, Ankara, Türkiye

Demirhan Diracoglu
Istanbul University, Istanbul Faculty of Medicine, Department of Physiotherapy and Rehabilitation, Istanbul, Türkiye

Dilsad Sindel
Istanbul University, Istanbul Faculty of Medicine, Department of Physical Therapy and Rehabilitation, Istanbul, Türkiye

Emine Dervis
Gaziosmanpasa Hospital, Clinic of Dermatology, Istanbul, Türkiye

Zafer Turkoglu
University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Dermatology, Istanbul, Türkiye

Nahide Onsun
Bezmialem Vakif University Faculty of Medicine, Department of Dermatology, Istanbul, Türkiye

Bulent Acunas
Istanbul University, Istanbul Faculty of Medicine, Department of Radiology, Interventional Radiology, Istanbul, Türkiye

Nuri Cagatay Cimsit
Marmara University Faculty of Medicine, Department of Radiology, Istanbul, Türkiye

Baris Bakir
Istanbul University, Istanbul Faculty of Medicine, Department of Radiology, Istanbul, Türkiye

Ozgur Sogut
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of Emergency Medicine, Istanbul, Türkiye

Zehra Zerrin Erkol
Abant Izzet Baysal University Faculty of Medicine, Department of Forensic Medicine, Bolu, Türkiye

The Medical Bulletin of Haseki

Scientific Advisory Board

Pelin Bagci
Marmara University Faculty of Medicine, Department of Pathology, Istanbul, Türkiye

Halide Nur Urer
University of Health Sciences Türkiye, Yedikule Chest Diseases and Thoracic Surgery Training and Research Hospital, Clinic of Pathology, Istanbul, Türkiye

Macit Koldas
University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital, Clinic of Medical Biochemistry, Istanbul, Türkiye

Alev Kural
University of Health Sciences Türkiye, Bakirkoy Dr. Sadi Konuk Training and Research Hospital, Clinic of Medical Biochemistry, Istanbul, Türkiye

Fikriye Uras
Marmara University Faculty of Pharmacy, Department of Medical Biochemistry, Istanbul, Türkiye

The Medical Bulletin of Haseki

Please refer to the journal's webpage **(https://www.hasekidergisi.com/**) for "About Us", "Instructions to Authors" and "Peer Review & Ethics".

The editorial and publication process of the Medical Bulletin of Haseki are shaped in accordance with the guidelines of ICMJE, WAME, CSE, COPE, EASE, and NISO. The journal is in conformity with the Principles of Transparency and Best Practice in Scholarly Publishing.

The Medical Bulletin of Haseki is indexed in **Emerging Sources Citation Index (ESCI), EBSCO Database, Gale, Turkish Medline-National Citation Index, Excerpta Medica/EMBASE, SCOPUS, TÜBİTAK/ULAKBİM, CINAHL, DOAJ, Hinari, GOALI, ARDI, OARE, AGORA, ProQuest, J-Gate, IdealOnline** and **Türkiye Citation Index**.

Owner: Mehmet Mesut Sonmez on University of Health Sciences Türkiye, Istanbul Haseki Training and Research Hospital

Responsible Manager: Akif ERBIN

The Medical Bulletin of Haseki

Contents

Review

362 **Current Advances in Genetics, Biomarkers, Neuroimaging, and Therapeutic Strategies in Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy: A Narrative Review**
Naga Bharathi Marni, Sowmya Karaka, Pranavi Badithala, Vinod Kumar Mugada; Andhra Pradesh, India

Original Articles

372 **The Impact of Nutritional Approach on Length of Stay and Costs in Coronary Intensive Care Patients: A Retrospective Analysis**
Cagla Ayer, Huseyin Demir, Tuncay Akar, Suleyman Mertoglu; Izmir, Mugla, Türkiye

384 **Association of SGLT-2 Inhibitor Use with Wound Healing and Amputation Risk in Patients with Diabetic Foot Ulcers: A Prospective Observational Study**
Zeynep Gizem Totik, Sevinc Puren Yucel Karakaya, Tayyibe Saler; Adana, Türkiye

391 **Anxiety, Depression, and Sleep Disturbances in Patients with Liver Cirrhosis: Associations with Disease Severity**
Zeynep Busra Akin, Mustafa Can Senoymak, Alpaslan Tanoglu, Suleyman Bas, Murat Yeniceri; Istanbul, Türkiye

399 **Sonographic Predictors of Malignancy in BI-RADS 4A Breast Lesions: A Diagnostic Performance Analysis**
Gulbanu Guner, Sevdiye Nur Agackaya, Fatma Kulali; Istanbul, Türkiye

407 **Sinoatrial Node Function Remains Unaltered During Transcutaneous Electrical Stimulation: An Electrophysiological Study**
Halime Kibar, Eser Kalaoglu, Ayse Karaoglu, Selim Sezikli, Tugba Aydin, Kemal Sitki Turker, Ilhan Karacan; Istanbul, Türkiye

415 **Temporal Muscle Thickness and Long-term Survival after Surgical Treatment of Chronic Subdural Hematoma: A 10-year Single-center Retrospective Cohort Study**
Bilal Bahadir Akbulut, Elif Ezgi Cenberlitas, Mustafa Serdar Boluk, Huseyin Biceroglu, Erkin Ozgiray, Taskin Yurtseven; Izmir, Türkiye

426 **Clinical Predictors of Proximal Esophageal Impaction and Their Implications for Management in Pediatric Foreign Body Ingestion: A Five-year Retrospective Cohort Study**
Mehmet Necmeddin Sutasir, Ahmet Bas, Betul Beyza Ozbal, Fatma Sarac, Banu Arslan; Istanbul, Türkiye

435 **Predictors of Stone-free Status Following Ultrathin Semi-rigid Ureteroscopy for Mid and Distal Ureteral Stones in Children**
Ayberk Iplikci, Muhammet Cicek, Ozgur Arikan, Mammad Khalilov, Mesrur Selcuk Silay, Asif Yildirim; Sinop, Malatya, Istanbul, Türkiye

442 **Guideline Concordance, Information Quality, and Readability of Large Language Model Responses to Female Urethral Stricture Questions: A Comparative Cross-sectional Study**
Cagri Sevik, Metin Savun, Suleyman Sahin, Yigit Can Filtekin, Halil Lutfi Canat; Istanbul, Türkiye

Case Reports

449 **Atypical Presentation of Primary Central Nervous System Lymphoma in a Young Adult: A Case Report and Current Literature Review**
Merve Dur Ince, Duygu Turkbey Simsek, Ferhat Cuce; Ankara, Türkiye

454 **Retrograde Bypass Due to Chronic Mesenteric Ischemia after Left Ventricular Aneurysm in a Patient with Woven Coronary Artery Anomaly: A Case Report and Current Literature Review**
Osman Fehmi Beyazal, Bilge Ecemis Yilmaz, Korhan Erkanli; Istanbul, Türkiye

The Medical Bulletin of Haseki

Contents

Letters to the Editor

458 **Distinguishing Sarcopenia From Parkinsonian Motor Dysfunction: Limitations of SARC-F Screening**
Okan Sokmen; Ankara, Türkiye

460 **The Importance of Structured Neuropathic Symptom Assessment in Kinesio Taping Studies on Carpal Tunnel Syndrome**
Taner Dandinoglu, Erkan Kaya; Bursa, Türkiye

462 **Beyond One-hour Outcomes in Pediatric Gastroenteritis Rehydration**
Emre Aygun; Istanbul, Türkiye

464 **Corrigendum**

2026 Reviewer Index

Review

DOI: 10.4274/haseki.galenos.2026.92486
Med Bull Haseki 2026;64(5):362-371

Current Advances in Genetics, Biomarkers, Neuroimaging, and Therapeutic Strategies in Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy: A Narrative Review

Naga Bharathi Marni, Sowmya Karaka, Pranavi Badithala, Vinod Kumar Mugada

Vignan Institute of Pharmaceutical Technology, Department of Pharmacy Practice, Visakhapatnam, Andhra Pradesh, India

Abstract

Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL), the most common monogenic small vessel disease, is caused by cysteine-altering mutations in the *neurogenic locus notch homolog protein 3* gene, which disrupt vascular smooth muscle cell function through protein misfolding and impaired vascular reactivity. Neuroimaging, particularly high-resolution magnetic resonance imaging, reveals characteristic white matter hyperintensities, and advanced imaging combined with deep learning enhances diagnostic precision. Blood neurofilament light chain is an emerging biomarker. Genetic counselling is crucial for early diagnosis and family screening. Current management involves vascular risk reduction, migraine control, and supportive care. Investigational therapies, including monoclonal antibodies, exon-skipping antisense oligonucleotides, CRISPR-based gene editing, and ongoing trials such as AMCAD and tocotrienol (NCT04658823), hold promise. Challenges remain in biomarker validation, therapeutic translation, and equitable access to diagnostics, highlighting the need for international collaboration to optimise CADASIL care.

Keywords: CADASIL, NOTCH3 mutation, biomarkers, neuroimaging, vascular dementia, small vessel disease

Introduction

Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL) is the most prevalent monogenic small vessel disease, caused by heterozygous cysteine-altering variants in the neurogenic locus notch homolog protein 3 (NOTCH3) extracellular domains *NOTCH3* gene (Figure 1). These mutations result in progressive cerebral small-vessel pathology, ultimately causing ischaemic strokes, leukoencephalopathy, and vascular dementia (1). Advances in whole-exome and genome sequencing have enabled genotype-phenotype correlations, notably linking Epidermal growth factor-like repeat (EGFr) domains 1-6 with early stroke onset and poor outcomes (2). Immunotherapies targeting aggregated NOTCH3 extracellular domains (NOTCH3ECD) have shown preclinical efficacy (3). High-resolution magnetic resonance imaging (MRI) enhances lesion characterisation, particularly white matter hyperintensities (WMHs) in the anterior temporal lobes and external capsules, which aid in diagnosis and tracking disease progression (4,5). Blood neurofilament light chain (NfL) is an emerging non-invasive biomarker correlating with imaging and clinical severity. This review highlights CADASIL's genetic basis, advances in sequencing, neuroimaging, and management, while underscoring emerging priorities: risk prediction, circulating/imaging biomarkers, NOTCH3ECD-dependent WMHs-and the role of international collaborative networks in translational research and clinical care (6).

Literature Search Strategy

This narrative review was based on a structured search of PubMed/MEDLINE, Scopus, Web of Science, and Google Scholar for studies published between January 2000 and

Corresponding Author: Naga Bharathi Marni, MD, Vignan Institute of Pharmaceutical Technology, Department of Pharmacy Practice, Visakhapatnam, Andhra Pradesh, India
E-mail: bharathimarni@gmail.com **ORCID:** orcid.org/0009-0001-3881-7302
Received: 28.08.2025 **Accepted:** 27.07.2026 **Publication Date:** 17.11.2026

Cite this article as: Bharathi Marni N, Karaka S, Badithala P, Kumar MV. Current advances in genetics, biomarkers, neuroimaging, and therapeutic strategies in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy: a narrative review. Med Bull Haseki. 2026;64(5):362-371

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

March 2026. Keywords included "CADASIL", "*NOTCH3*", "genetics", "biomarkers", "MRI", "neuroimaging", "gene-therapy", "CRISPR", "vascular dementia", and "small vessel disease". The search identified about 1,240 records; after removing duplicates, 980 were screened by title and abstract. Of these, 660 full texts were assessed, and 410 were excluded for irrelevance, non-English language, or lack of full text. A total of 250 studies were included, comprising original research, systematic reviews, meta-analyses, clinical trials, consensus statements, and clinical guidelines. Table 1 summarises selected CADASIL studies by study design, population or experimental model, and key findings.

Epidemiology and Natural History

Prevalence and Demographic Distribution

Genomic analyses using large databases such as ExAC and UK Biobank have significantly revised CADASIL prevalence estimates, identifying pathogenic *NOTCH3* variants in up to one in three hundred individualssubstantially higher than earlier registry-based estimates (7) (Figure 1). Regional variation is evident, with the p.Arg544Cys variant found in 0.85% of the Taiwanese population. Demographic factors influence clinical presentation. Asian patients tend to have a later symptom onset (mid-50s) and a lower migraine prevalence (2.5-33% vs. 40-54%) compared to Western cohorts (mid-30s). Intracerebral haemorrhage and cerebral microbleeds are more common in Asian patients (25% and 87.5%, respectively) than in European patients (<1% and 31%, respectively). Conversely, anterior temporal pole WMHs-a hallmark feature-are less frequently observed in Asians (20-44.5%) than in Europeans (89-93%) (8).

Disease Progression and Clinical Milestones

CADASIL severity is measured by the NOTCH3-SVD Staging System, integrating Fazekas' scores, lacune count, and mRS, with 94% showing sequential progression (9), and the Clinical CADASIL Grading System, demonstrating high inter-rater reliability (κ=0.855) (10). Long-term studies reveal 85% progression over 18 years, correlating more strongly with imaging and biological markers-lacune burden, parenchymal loss, and NOTCH3 aggregation-than with clinical status, age, or sex (1,9,11).

Genetic Basis of CADASIL and Clinical Genetic Counselling Guidelines

CADASIL arises from cysteine-altering mutations within the 34 EGFr repeats of the *NOTCH3* gene (1). These disrupt disulphide bonding, causing extracellular domain misfolding, aberrant bridge formation, aggregation, and ER retention. Misfolded NOTCH3ECD accumulates as granular osmiophilic material (GOM) in VSMCs and pericytes, sequestering TIMP3, clusterin, and LTBP-1, disrupting perivascular homeostasis (12,13). This impairs cerebrovascular reactivity and drainage and promotes degeneration, leading to hypoperfusion, white-matter lesions, and vascular dementia (14). Pathogenic variants include P.R544C and p.Arg133Cys, with EGFr 7-34 linked to broader small vessel disease phenotypes (15-17).

Genetic Testing Indications

Genetic testing for *NOTCH3* is recommended when clinical, imaging, and familial features suggest CADASIL, including early-onset (<55 years) ischemic events, recurrent migraine with aura (Figure 2), confluent leukoencephalopathy in anterior temporal poles/external capsules, and a family history of stroke or cognitive decline. Confirmation requires identifying cysteine-altering EGFr variants via DNA sequencing (Figure 1). VUS are assessed using ACMG/AMP criteria (PS4, PP4, PM2, PS3) (18), refined by ClinGen guidelines emphasising cysteine-altering EGFr variants (Figure 1), with further resolution through segregation, computational, and domain-based stratification.

Biomarkers and Molecular Pathways

Emerging Molecular Pathways

Proteomic analyses in CADASIL vasculature reveal impaired ECM organisation, disrupted VSMC cytoskeletal dynamics, and defective proteostasis, promoting aggregation of mutant NOTCH3ECD (Figure 1). This drives autophagic dysfunction, dysregulated angiogenesis, and TGF-β signalling, leading to fibrotic microvascular remodelling (19). Mechanistically, N3ECD accumulation induces oxidative stress via Nox5, activating ER stress and RhoA/Rho kinase pathways, while disrupted calcium homeostasis impairs eNOS-sGC-cGMP signalling, compromising endothelial integrity (20,21). Targeted therapies-antisense oligonucleotide (ASO) for exon skipping, monoclonal antibodies against NOTCH3ECD, RNA interference, and CRISPR-mediated editing-show promise in reducing aggregates, restoring vascular function, and correcting pathogenic alleles in preclinical models (14).

Fluid and Imaging Biomarkers

Neurofilament light chain in blood is a validated biomarker of neuroaxonal injury in CADASIL. Elevated levels correlate with white matter lesion burden, cognitive decline, and long-term disability and predict outcomes over 7-17 years, supporting its use in disease monitoring. In contrast, although microRNAs have been studied in related cerebrovascular disorders, they have not been clinically validated for CADASIL. Retinal imaging detects early microvascular changes in CADASIL, manifested as reduced vascular density on optical coherence

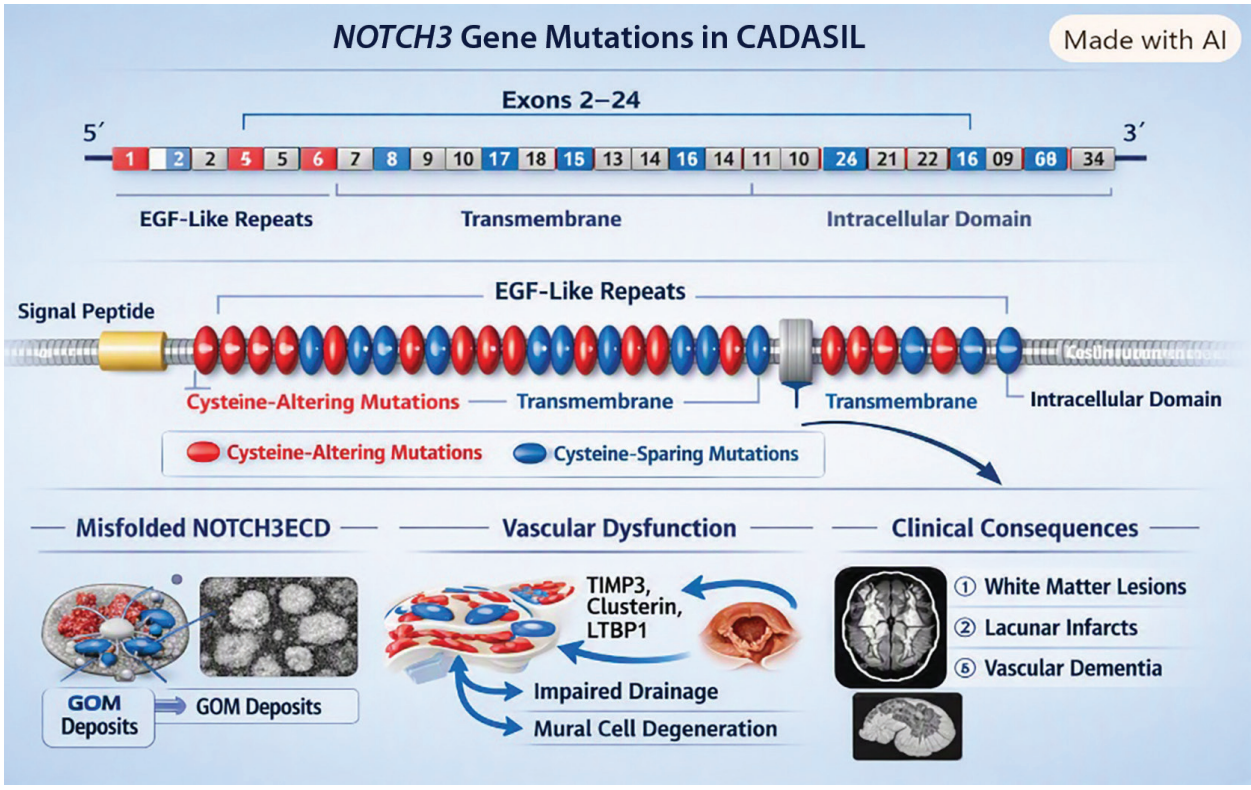


Figure 1. Schematic representation of the *NOTCH3* gene and cysteine dependent mutations in CADASIL. The *NOTCH3* gene comprises 33 exons encoding multiple domains, with most pathogenic variants located within the epidermal growth factor (EGF) like repeats (exons 2–24). Cysteine altering mutations (red) disrupt disulphide bonding, leading to misfolded NOTCH3 extracellular domain, granular osmiophilic material (GOM) deposition, and vascular dysfunction. Cysteine sparing mutations (blue) are less pathogenic. The schematic illustrates the relationship between molecular alterations, impaired extracellular matrix organization, and clinical consequences including white matter lesions and vascular dementia (21)

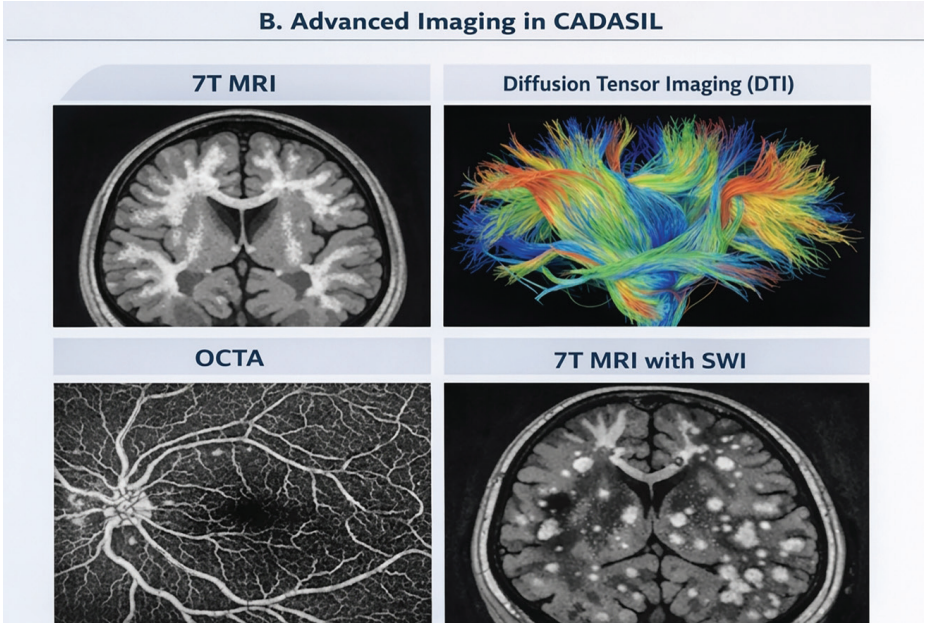


Figure 2. Clinical manifestations and MRI features of CADASIL, showing common symptoms (small subcortical infarcts, migraines with aura, cognitive impairment, psychiatric disturbances) and hallmark imaging findings (white matter hyperintensities, lacunes, cerebral microbleeds, enlarged perivascular spaces) in characteristic brain regions (21)

tomography angiography (OCTA) and an increased wall-to-lumen ratio on adaptive optics. However, these alterations do not correlate with MRI-defined lesion burden or clinical severity, limiting their prognostic and diagnostic applicability (22,23).

Challenges in Ensuring Consistent and Reproducible Results for Biomarker Measurement

The reproducibility of biomarker findings in CADASIL is limited by small cohort sizes, single-centre designs, and methodological inconsistencies in sample handling, assay platforms, and analytical protocols. While high-throughput techniques like mass spectrometry and multiplex immunoassays hold potential, they require rigorous standardisation and transparent performance reporting to be clinically useful (24). Multicentre collaborations with harmonised methodologies and open data sharing are essential for advancing biomarker validation in CADASIL.

Advances in Diagnostic Modalities

MRI and Advanced Neuroimaging

In CADASIL, hallmark WMHs occur bilaterally in anterior temporal poles and external capsules, with high diagnostic specificity even in early stages (1,25,26). WMHs also show symmetrical periventricular and deep white matter distribution, sometimes extending into the corpus callosum. Lacunar infarcts in basal ganglia, thalamus, centrum semiovale, and pons correlate with cognitive decline (1,27).

Temporal pole WMHs differ pathologically, reflecting enlarged perivascular spaces and myelin loss (28). Quantitative tools-CADA-MRIT and CADASILSeverity Grading System-standardise lesion assessment, enhance diagnostic accuracy and improve trial stratification (Figure 2) (10).

Limitations of Conventional MRI in CADASIL

Despite its diagnostic utility, conventional MRI has limited prognostic sensitivity in CADASIL; imaging markers such as WMHs and lacunes correlate poorly with clinical severity and do not reliably predict disease progression. Furthermore, standard MRI lacks the spatial resolution necessary to evaluate microvascular hemodynamics or to detect early perfusion changes, limiting its capacity to capture preclinical disease activity (29).

Quantitative and High-Field Techniques

Quantitative neuroimaging enhances CADASIL assessment. SIENA-derived PBVC shows ~0.45% annual atrophy, with incident lacunes predicting decline better than WMH progression (29). Diffusion tensor imaging reveals increased mean diffusivity and reduced peak height in normal-appearing white matter, enabling early detection (30). Ultra-high-field 7T SWI visualises microbleeds and vascular changes, including the "Chocolate Chip Sign" and 14.6% venous density reduction (31). Deep learning CNNs improve lesion segmentation for WMHs, lacunes, microbleeds, and perivascular spaces (Figure 2), with Dice ~0.75-0.85 (32). Limitations include high cost, infrastructure demands, implant incompatibility, methodological variability, and lack of clinician training (33,34).

Retinal and Extracerebral Imaging Windows

Retinal vascular changes-arterial narrowing, arteriovenous nicking, and reduced OCTA vessel density-may indicate early CADASIL pathology but correlate poorly with symptoms or MRI lesion burden (22,23,35). Reliability is limited by artefacts, segmentation errors, and protocol variability, especially in OCTA (35). Standardised acquisition, quality control, and analysis protocols are urgently needed for research and clinical use. Table 2 compares conventional MRI, diffusion tensor imaging, 7-T MRI, and optical coherence tomography angiography in terms of what each technique detects and its main limitations.

Laboratory Biomarkers

Genetic testing for pathogenic *NOTCH3* mutations remains the gold-standard assay for CADASIL, performed on blood DNA (36). No protein- or metabolite-based biomarker has achieved validation; CSF amyloid-β42, total tau, and phosphorylated tau lack specificity and are not recommended. Neurofilament light chain in blood and CSF correlates with severity, lesion burden, cognition, and outcomes, supporting monitoring utility (9,37). Emerging biomarkers include CNS-derived extracellular vesicles (EVs) carrying disease-specific cargo (38), though clinical use is limited by isolation variability, purity, and specificity; omics-based EV profiling is under investigation (19).

Current Management Strategies

General Principles of Care

Effective management of CADASIL centres on intensive blood pressure control and smoking cessation-interventions strongly associated with reduced risk of stroke and cognitive decline. Although evidence for treating comorbid dyslipidemia remains limited, lifestyle modifications and comprehensive vascular risk control contribute to slowing disease progression (39). Pharmacologic therapies, particularly antiplatelet and anticoagulant agents, require individualised risk-benefit assessments due to the elevated risk of intracerebral haemorrhage; routine use of antiplatelet agents is not recommended in patients without prior ischemic events (40). Table 3 summarises clinical recommendations for confirming the diagnosis, counselling at-risk families,

monitoring disease progression, and managing vascular risk and symptoms.

Antithrombotic and Anticoagulant Therapy

Evidence on antithrombotic therapy in CADASIL is limited. In a TriNetX cohort (n=455), antiplatelet use did not alter ischemic stroke or ICH risk (41). A 2024 review of 142 ICH cases found anticoagulants-not antiplatelets-associated with higher hemorrhagic risk (20.0% vs. 1.9%, p=0.006), with ICH in 5.9% overall and 6.7% linked to anticoagulants (40,42). Low-dose aspirin after ischemic events appears generally safe (39), though primary prevention remains uncertain. Anticoagulation is reserved for atrial fibrillation (CHA_2DS_2-VASc ≥2) or venous thromboembolism, with alternatives like left atrial appendage occlusion considered in patients with extensive microbleeds.

Acute Reperfusion Therapies

Evidence for acute reperfusion in CADASIL is limited. Individualised intravenous thrombolysis with alteplase has shown favourable outcomes; in a multicentre series (n=12), 83% achieved mRS 0-1 at 90 days without symptomatic haemorrhage (43,44). Consensus

Table 1. Selected studies and recommendations on the epidemiology, genetics, diagnosis, biomarkers, and management of CADASIL

Author	Year	Study type	Sample/model	Key findings
Yamamoto et al. (1)	2023	Review	NA	Updated epidemiology, biomarkers, and pathogenesis of CADASIL
Choi (2)	2010	Review	NA	Genetic basis of CADASIL and *NOTCH3* mutations
Oliveira et al. (3)	2022	Experimental	Mouse model	Immunotherapy reduces NOTCH3 aggregation
Rutten et al. (7)	2020	Cohort	UK Biobank (>1000)	High prevalence and broad phenotype of *NOTCH3* variants
Kim et al. (8)	2022	Observational	Asian cohort	Ethnic differences in phenotype and imaging features
Gravesteijn et al. (9)	2025	Cohort	1900+	Disease severity staging system (NOTCH3-SVD)
Anisetti et al. (10)	2023	Cross-sectional	Multicenter	CADASIL clinical grading system validation
Alirezaei et al. (37)	2019	Review	NA	NfL correlation with MRI and disease severity
Meschia et al. (39)	2023	Guideline	AHA statement	Management recommendations for CADASIL

CADASIL: Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy, *NOTCH3*: Neurogenic locus notch homolog protein 3, MRI: Magnetic resonance imaging

Table 2. Advantages and limitations of imaging modalities used in CADASIL

Imaging Method	Advantage	Limitation
Conventional MRI	High sensitivity for lesion detection	Poor prognostic value
Diffusion tensor imaging (DTI)	Detects microstructural changes	Protocol variability
7 Tesla MRI (7T MRI)	High-resolution vascular imaging	Expensive and limited availability
Optical coherence tomography angiography (OCTA)	Early microvascular detection	Poor correlation with clinical severity

MRI: Magnetic resonance imaging

Table 3. Clinical recommendations for the diagnosis, management, and genetic counselling of patients with CADASIL

Clinical domain	Current recommendation
Blood pressure management	Maintain blood pressure <130/80 mmHg through intensive vascular risk control
Smoking cessation	Complete smoking cessation is strongly recommended
Genetic diagnosis	Confirm diagnosis by identifying pathogenic *NOTCH3* mutations using genetic testing
Genetic counselling	Offer pre-test and post-test genetic counselling and cascade testing for at-risk relatives
Neuroimaging	Brain MRI is recommended for diagnosis and monitoring disease progression
Antiplatelet therapy	Consider after ischemic stroke or TIA; not routinely recommended for primary prevention
Anticoagulation	Use only for standard indications (e.g., atrial fibrillation) after individualized risk-benefit assessment
Symptomatic management	Appropriate treatment of migraine, cognitive impairment, psychiatric symptoms, and rehabilitation should be individualized

NOTCH3: Neurogenic locus notch homolog protein 3, MRI: Magnetic resonance imaging

discourages thrombolysis for lacunar strokes but supports cautious use in large-vessel occlusions, contingent on MRI microbleed absence and close monitoring. Mechanical thrombectomy has shown promise in rare large-vessel occlusions, with successful reperfusion yielding clinical improvement without haemorrhagic complications (45).

Symptomatic and Supportive Treatments

Migraine Management in CADASIL

Acute migraine in CADASIL is managed with analgesics (paracetamol, ibuprofen, aspirin) and abortive agents like sumatriptan, which-despite vasoconstrictive concerns-has not been clearly linked to ischemic events and is considered safe short-term with individualised risk assessment (40,46). Prophylaxis includes amitriptyline, pizotifen, acetazolamide, and sodium valproate, while propranolol is avoided due to adverse effects (9,40). Emerging CGRP-targeted therapies (gepants, ditans) may be less vasoconstrictive but require caution until CADASIL safety is established (46,47).

Cognitive, Psychiatric, and Rehabilitative Therapies

Evidence supporting targeted interventions for neuropsychiatric symptoms in CADASIL remains sparse, with most treatment regimens adapted from general psychiatric practice rather than CADASIL-specific trials. The American Heart Association notes that optimal strategies for these symptoms have yet to be clearly defined. Nonetheless, selective serotonin reuptake inhibitors are frequently prescribed in clinical settings, although their efficacy in CADASIL has not been systematically validated (13).

Pharmacological Management of Neuropsychiatric Symptoms

Depression

Selective serotonin reuptake inhibitors are routinely used and generally well tolerated in treating depressive symptoms in CADASIL, with dose escalation considered for suboptimal responders. Although amitriptyline has demonstrated limited benefit, its use is discouraged in patients with cognitive impairment because of its anticholinergic effects. Sodium valproate may theoretically exert benefits via modulation of NOTCH3 receptor activity; however, mood disturbances in CADASIL are frequently resistant to pharmacotherapy. In rare, treatment-refractory melancholic cases, electroconvulsive therapy has been employed as a last-line intervention (48).

Psychosis

Short-term administration of antipsychotics, particularly atypical agents, has been used to manage psychotic symptoms such as agitation, aggression, and excitatory states. In one reported case, a combination of lithium and a high-potency typical antipsychotic was effective in treating mixed manic symptoms (48).

Cognitive Dysfunction

No pharmacological agents have consistently improved executive dysfunction in CADASIL. A randomised controlled trial of donepezil showed no significant impact on global cognition but yielded modest executive-function domains at a 10-mg/day dose. Galantamine has shown limited benefit in isolated cases of CADASIL-related dementia, which supports the cholinergic deficit hypothesis; however, robust evidence from larger trials is still lacking (48).

Psychosocial Interventions in CADASIL Management

Family-Centred Support

Tailored psychotherapeutic interventions addressing the psychosocial challenges of hereditary cerebrovascular disease should be systematically offered to CADASIL patients and their family members to provide emotional and psychological support throughout the disease course (48).

Comprehensive Care

While psychotropic medications may be employed under clinical guidelines, their use requires vigilant monitoring of blood pressure and cognitive status. Despite growing consensus on the value of multidisciplinary care-including cognitive, physical, and psychosocial rehabilitation-robust evidence comparing the effectiveness of these approaches remains limited. A recent case report documented that structured physical therapy resulted in full recovery of limb strength, although aphasia and cognitive impairments persisted, highlighting the need for individualised, symptom-specific interventions and sustained follow-up to optimise functional outcomes (40,48).

Emerging and Future Therapies in Cadasil: Molecular and Translational Perspectives

Targeted Molecular Therapies

Emerging strategies for CADASIL aim to counteract the pathogenic consequences of *NOTCH3* mutations in vascular smooth muscle cells by restoring proper gene and protein function. Genetic correction techniques such as ASO therapy and CRISPR/Cas9-mediated exon deletion have demonstrated success in restoring exon integrity and normal NOTCH3 processing in vitro and in patient-derived cells. More recently, a dual adeno-associated virus (AAV)-split ABEmax base-editing platform achieved up to 8.82% correction of the R421C mutation in vascular organoids derived from patient cells (49).

Future Directions in Clinical Management

Current research focuses on developing reliable blood-based biomarkers, refining diagnostic criteria,

and conducting longitudinal studies to map disease progression. Therapeutic innovations are also underway, including gene silencing, exon skipping, and monoclonal antibodies targeting pathogenic NOTCH3 mechanisms. Looking ahead, gene-editing technologies and RNA-based therapies offer promising avenues for targeted disease modification (50,51).

Immunotherapy Approaches

Both active vaccination and passive immunotherapy targeting the aggregated NOTCH3 extracellular domain (ECD) have shown efficacy in transgenic mouse models. Active immunisation reduced perivascular NOTCH3 deposition by up to 48%, while passive monoclonal antibodies, including 5E1 and A13, improved cerebrovascular reactivity, reduced mural cell loss and modulated biomarkers in mutant models (3,50,51).

Additional Experimental Therapies

Further investigational avenues include gamma-secretase inhibitors (e.g., MRK-003) to regulate NOTCH signalling (12,13), molecular chaperones such as the BRICHOS domain to prevent aggregation (52). These approaches predominantly act by disrupting NOTCH3 extracellular domain aggregation and preserving vascular function. Table 4 presents current approaches to CADASIL care alongside investigational strategies, including gene editing, anti-NOTCH3 antibodies, and RNA-based therapies, and indicates their stage of development.

Challenges in Clinical Translation

Clinical translation of NOTCH3-targeted therapies remains constrained by an incomplete understanding of CADASIL pathophysiology and a lack of validated therapeutic endpoints, and limited evidence linking GOM or NOTCH3 extracellular domain deposition to clinical progression. Additional obstacles include the difficulty of purifying pathological aggregates and the inadequacy of existing animal models to fully recapitulate the human disease. The development of mutation-specific biomarkers and predictive assays is urgently required to guide therapeutic evaluation across CADASIL genotypes (13).

Clinical Trials, Collaborative Frameworks, and Research Priorities in CADASIL

Ongoing Clinical Trials

A multicentre phase II trial, AMCAD, is evaluating intravenous adrenomedullin (15 ng/kg/min for 8 hours daily over 14 days) in 60 genetically confirmed CADASIL patients, with changes in frontal cortex cerebral blood flow at day 28 measured via arterial spin labelling MRI as the primary endpoint (53). Concurrently, a randomised, placebo-controlled phase II study (NCT04658823) is assessing palm-derived tocotrienol supplementation over 24 weeks to evaluate its potential to slow progression of WMHs by modulating oxidative and inflammatory pathways.

Quantitative Biomarkers in Research

Biomarkers such as the quantitative NOTCH3 score, which measures granular deposition surface area in arterioles, and serum NfL, which correlates with infarct volume and remains elevated post-stroke, enhance sensitivity, enable earlier efficacy assessments, and improve patient stratification in both preclinical and clinical studies (7).

Collaborative Research Frameworks

To address the challenges posed by CADASIL's rarity, global collaborations are critical. The Global CADASIL Consortium, involving 46 investigators from 17 countries, has harmonised data collection across 14 cohorts (24). In Europe, the ERN-RND registry, implemented via Research Electronic Data Capture and the Clinical Patient Management System, standardises clinical and biomarker data from 24 centres in accordance with the Findable, Accessible, Interoperable, Reusable principles. Within this framework, the Solve-RD initiative integrates multi-omics data and phenotype ontologies to enable reanalysis of undiagnosed cases and to improve cross-centre interpretive consistency.

Table 4. Current management approaches and emerging therapeutic strategies for CADASIL

Current standard therapy	Emerging therapeutic strategy	Current status
Blood pressure control and vascular risk reduction	CRISPR/Cas9 gene editing	Preclinical
Smoking cessation and lifestyle modification	AAV-mediated base editing	Experimental
Migraine management	Anti-NOTCH3 monoclonal antibodies	Animal studies
Supportive care	Antisense oligonucleotide (ASO) therapy	Preclinical
Conventional symptomatic treatment	RNA interference (RNAi)	Experimental
Rehabilitation and multidisciplinary care	Molecular chaperones (BRICHOS)	Experimental
Conventional clinical management	Precision medicine using biomarkers and AI-assisted imaging	Future direction

AAV: Adeno-associated virus, *NOTCH3*: Neurogenic locus notch homolog protein 3, AI: Artificial intelligence

Key Recommendations

Clinical Management Priorities

Strict vascular risk control, especially blood pressure optimisation (<130/80 mmHg) and smoking cessation. Routine counselling/psychological assessments to detect cognitive and mood changes early. Proactive management of migraine and monitoring tools (NOTCH3-SVD, symptoms). Family-based genetic counselling to inform at-risk relatives. Use of biomarker-guided staging (NfL, CNS-derived EVs) with validated imaging tools (NOTCH3-SVD, CADASIL-MRIT) for monitoring and research readiness.

Research Imperatives

Conduct longitudinal, multiethnic cohort studies to clarify genotype–phenotype correlations. Validate emerging biomarkers for diagnosis and monitoring. Advance therapeutic candidates-particularly NOTCH3-extracellular domain-targeting antibodies (5E1,A13) and gene-editing tools (CRISPR/Cas9, base editors)-towards clinical application. Expand artificial intelligence-based lesion quantification using federated learning to enhance comparability across centres while preserving data security.

Global Collaborative Goals

Harmonise *NOTCH3* genetic screening protocols internationally. Develop ethically sound predictive testing frameworks. Implement adaptive, biomarker-stratified Bayesian clinical trial designs. Establish unified standards for diagnostics, surrogate endpoints, and mechanism-driven interventions. Promote secure, equitable global data sharing to accelerate therapeutic innovation in rare cerebrovascular diseases.

Limitations of the Review

As a narrative review, this article has limitations, including the absence of a formal systematic search strategy with risk-of-bias assessment and the potential for selection bias in the inclusion of studies. Efforts were made to include recent, high-quality, and clinically relevant evidence to provide a comprehensive overview of advances in CADASIL.

Conclusion

CADASIL research is transitioning from symptomatic care towards precision medicine. Advances in genomics, circulating biomarkers, quantitative neuroimaging, and AI-assisted image analysis are refining diagnosis and monitoring. Emerging molecular therapies-ASOs, monoclonal antibodies targeting NOTCH3, RNA-based therapeutics, and CRISPR gene editing-offer genuine disease-modifying potential, though still experimental. This review underscores that integrating genomic profiling, validated biomarkers, advanced imaging, and precision therapeutics will reshape CADASIL management, enabling earlier diagnosis, personalised intervention, and improved long-term outcomes. The main contribution lies in synthesising evidence across molecular, imaging, and therapeutic domains, highlighting how international collaboration, biomarker standardisation, and multicentre trials are essential to accelerate translation into clinical practice and establish CADASIL as a model for precision neurology.

Footnotes

Authorship Contributions

Concept: N.B.M., S.K., P.B., V.K.M., Design: N.B.M., S.K., P.B., V.K.M., Data Collection or Processing: N.B.M., S.K., P.B., V.K.M., Analysis or Interpretation: N.B.M., S.K., P.B., V.K.M., Literature Search: N.B.M., S.K., P.B., V.K.M., Writing: N.B.M., S.K., P.B., V.K.M.

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

References

1. Yamamoto Y, Liao YC, Lee YC, Ihara M, Choi JC. Update on the epidemiology, pathogenesis, and biomarkers of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. J Clin Neurol. 2023;19:12-27.
2. Choi JC. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy: a genetic cause of cerebral small vessel disease. J Clin Neurol. 2010;6:1-9.
3. Oliveira DV, Coupland KG, Shao W, et al. Active immunotherapy reduces NOTCH3 deposition in brain capillaries in a CADASIL mouse model. EMBO Mol Med. 2023;15:e16556.
4. Goldstein ED, Majersik JJ, McNally S. Intracranial high-resolution vessel wall imaging in CADASIL. Neurology. 2020;94:1040-1.
5. Taniguchi A, Shindo A, Tabei KI, et al. Imaging characteristics for predicting cognitive impairment in patients with cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. Front Aging Neurosci. 2022;14:876437.
6. Kerner C, Fadhil A, Sundararajan S. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy: a woman with lacunar stroke. Stroke. 2022;53:e162-4.
7. Rutten JW, Hack RJ, Duering M, et al. Broad phenotype of cysteine-altering NOTCH3 variants in UK Biobank: CADASIL to nonpenetrance. Neurology. 2020;95:e1835-43.
8. Kim Y, Bae JS, Lee JY, et al. Genotype and phenotype differences in CADASIL from an Asian perspective. Int J Mol Sci. 2022;23:11506.
9. Gravesteijn G, Rutten JW, Cerfontaine MN, et al. Disease severity staging system for NOTCH3-associated small vessel disease, including CADASIL. JAMA Neurol. 2025;82:49-60.
10. Anisetti B, Greco E, Stojadinovic E, et al. Novel grading system for CADASIL severity: a multicenter cross-sectional study. Cereb Circ Cogn Behav. 2023;5:100170.

11. Cerfontaine MN, Gravesteijn G, Hack RJ, et al. Association between vascular NOTCH3 aggregation and disease severity in a CADASIL cohort-implications for NOTCH3 variant-specific disease prediction. Ann Neurol. 2025;98:273-85.
12. Schoemaker D, Arboleda-Velasquez JF. Notch3 signaling and aggregation as targets for the treatment of CADASIL and other NOTCH3-associated small-vessel diseases. Am J Pathol. 2021;191:1856-70.
13. Locatelli M, Padovani A, Pezzini A. Pathophysiological mechanisms and potential therapeutic targets in cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL). Front Pharmacol. 2020;11:321.
14. Young KZ, Ramírez CR, Keep SG, et al. Oligomerization, trans-reduction, and instability of mutant NOTCH3 in inherited vascular dementia. Commun Biol. 2022;5:3259.
15. Mönkäre S, Kuuluvainen L, Schleutker J, et al**.** Clinical features and spectrum of NOTCH3 variants in Finnish patients with cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL). Acta Neurol Scand. 2022;146:643-51.
16. Hu Y, Sun Q, Zhou Y, et al. NOTCH3 variants and genotype-phenotype features in Chinese CADASIL patients. Front Genet. 2021;12:705284.
17. Cheng Y, Liao Y, Chen C, et al. Contribution of the APOE genotype to cognitive impairment in individuals with NOTCH3 cysteine-altering variants. J Am Heart Assoc. 2023;12:e032689.
18. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405-24.
19. Muiño E, Fernández-Cadenas I, Arboix A. Contribution of "omic" studies to the understanding of CADASIL: a systematic review. Int J Mol Sci. 2021;22:7357.
20. Heidari P, Taghizadeh M, Vakili O. Signaling pathways and molecular mechanisms involved in the onset and progression of cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy (CADASIL); a focus on Notch3 signaling. J Headache Pain. 2025;26:20.
21. Neves K, Morris HE, Lopes RA, et al. Peripheral arteriopathy caused by Notch3 gain-of-function mutation involves ER and oxidative stress and blunting of NO/sGC/cGMP pathway. Clin Sci (Lond). 2021;135:753-73.
22. Krivosic V, Paques M, Hervé D, et al. Retinal vascular density in CADASIL (cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy). BMJ Neurol Open. 2023;5:e000417.
23. Paques M, Krivosic V, Castro-Farias D, et al. Early remodeling and loss of light-induced dilation of retinal small arteries in CADASIL. J Cereb Blood Flow Metab. 2024;44:1089-101.
24. Saks DG, Bajorek B, Catts VS, et al. The protocol for an observational Australian cohort study of CADASIL: the AusCADASIL study. Cereb Circ Cogn Behav. 2024;6:100225.
25. O'Sullivan M, Jarosz JM, Martin RJ, et al. MRI hyperintensities of the temporal lobe and external capsule in patients with CADASIL. Neurology. 2001;56:628-34.
26. Markus HS, Martin RJ, Simpson MA, et al. Diagnostic strategies in CADASIL. Neurology. 2002;59:1134-8.
27. Jolly AA, Nannoni S, Edwards H, et al. Prevalence and predictors of vascular cognitive impairment in patients with CADASIL. Neurology. 2022;99:e453-61.
28. Trikamji B, Thomas M, Hathout G, et al. An unusual case of cerebral autosomal-dominant arteriopathy with subcortical infarcts and leukoencephalopathy with occipital lobe involvement. Ann Indian Acad Neurol. 2016;19:272.
29. Ling Y, Guio FD, Jouvent E, et al. Clinical correlates of longitudinal MRI changes in CADASIL. J Cereb Blood Flow Metab. 2018;39:1299-305.
30. Jouvent E, Duering M, Chabriat H. Cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. Stroke. 2020;51:21-8.
31. Ando S, Saito R, Kitahara S, et al. "Chocolate chip sign" on susceptibility-weighted imaging. Neurol Genet. 2025;11:e200237.
32. Demeusy V, Roche F, Vincent F, et al. Development and validation of a two-stage convolutional neural network algorithm for segmentation of MRI white matter hyperintensities for longitudinal studies in CADASIL. Comput Biol Med. 2024;180:108936.
33. Chen X, Qu L, Xie Y, et al. A paired dataset of T1- and T2-weighted MRI at 3 Tesla and 7 Tesla. Sci Data. 2023;10:24.
34. de Vries E, Hagbohm C, Ouellette R, et al. Clinical 7 Tesla magnetic resonance imaging: impact and patient value in neurological disorders. J Intern Med. 2025;297:244-61.
35. Biffi E, Turple Z, Chung J, et al. Retinal biomarkers of cerebral small vessel disease: a systematic review. PLoS One. 2022;17:e0266974.
36. Park JY, Song JY, Chang JY, et al. Genetic and imaging features of CADASIL patients with acute ischemic stroke. Sci Rep. 2025;15:2201.
37. Alirezaei Z, Pourhanifeh MH, Borran S, et al. Neurofilament light chain as a biomarker, and correlation with magnetic resonance imaging in diagnosis of CNS-related disorders. Mol Neurobiol. 2019;57:469-91.
38. Wang X, Yang H, Liu C, et al. A new diagnostic tool for brain disorders: extracellular vesicles derived from neuron, astrocyte, and oligodendrocyte. Front Mol Neurosci. 2023;16:1194210.
39. Meschia JF, Worrall BB, Elahi FM, et al. Management of inherited CNS small vessel diseases: the CADASIL example: a scientific statement from the American Heart Association. Stroke. 2023;54:e286-303.
40. Servito M, Gill I, Durbin J, et al. Management of coronary artery disease in CADASIL patients: review of current literature. Medicina. 2023;59:586.
41. Muppa J, Yaghi S, Goldstein ED. Antiplatelet use and CADASIL: a retrospective observational analysis. Neurol Sci. 2023;44:2831-4.

42. Sukhonpanich N, Markus HS. Prevalence, clinical characteristics, and risk factors of intracerebral haemorrhage in CADASIL: a case series and systematic review. J Neurol. 2024;271:2423-33.

43. Pescini F, Torricelli S, Squitieri M, et al. Intravenous thrombolysis in CADASIL: report of two cases and a systematic review. Neurol Sci. 2022;44:491-8.

44. Rinnoci V, Nannucci S, Valenti R, et al. Cerebral hemorrhages in CADASIL: report of four cases and a brief review. J Neurol Sci. 2013;330:45-51.

45. Xiao S, Ke M, Cai K, et al**.** Treatment options for patients with CADASIL and large-scale cerebral infarction: mechanical thrombectomy and antiplatelet therapy-a case report. Front Neurol. 2024;15:1400537.

46. de Boer I, MaassenVanDenBrink A, Terwindt GM. The potential danger of blocking CGRP for treating migraine in CADASIL patients. Cephalalgia. 2020;40:1676-8.

47. de Boer I, Verhagen IE, Souza M, et al. Place of next generation acute migraine specific treatments among triptans, non-responders and contraindications to triptans and possible combination therapies. Cephalalgia. 2023;43:03331024221143773.

48. Chabriat H, Lesnik Oberstein S. Cognition, mood and behavior in CADASIL. Cereb Circ Cogn Behav. 2022;3:100043.

49. Wang J, Zhang L, Wu G, et al. Correction of a CADASIL point mutation using adenine base editors in hiPSCs and blood vessel organoids. J Genet Genomics. 2024;51:197-207.

50. Ghezali L, Capone C, Baron-Menguy C, et al. Notch3ECD immunotherapy improves cerebrovascular responses in CADASIL mice. Ann Neurol. 2018;84:246-59.

51. Machuca-Parra AI, Bigger-Allen A, Sanchez AV, et al. Therapeutic antibody targeting of Notch3 signaling prevents mural cell loss in CADASIL. J Exp Med. 2017;214:2271-82.

52. Oliveira DV, Svensson J, Zhong X, et al. Molecular chaperone BRICHOS inhibits CADASIL-mutated NOTCH3 aggregation in vitro. Front Mol Biosci. 2022;9:812808.

53. Washida K, Saito S, Tanaka T, et al. A multicenter, single-arm, phase II clinical trial of adrenomedullin in patients with cerebral autosomal dominant arteriopathy with subcortical infarcts and leukoencephalopathy. Cereb Circ Cogn Behav. 2024;6:100211.

Original Article

DOI: 10.4274/haseki.galenos.2026.74436
Med Bull Haseki 2026;64(5):372-383

The Impact of Nutritional Approach on Length of Stay and Costs in Coronary Intensive Care Patients: A Retrospective Analysis

Cagla Ayer[1], Huseyin Demir[2], Tuncay Akar[3], Suleyman Mertoglu[4]

[1]*Izmir Katip Celebi University Faculty of Health Sciences, Department of Nutrition and Dietetics, Izmir, Türkiye*
[2]*Izmir Katip Celebi University Faculty of Economics and Administrative Sciences, Department of Health Management, Izmir, Türkiye*
[3]*Mugla Sitki Kocman University, Institute of Health Sciences, Department of Health Management, Mugla, Türkiye*
[4]*Izmir Provincial Directorate of Health, Izmir, Türkiye*

Abstract

Aim: Early and appropriate nutritional support plays a fundamental role in the management of critically ill patients; however, its clinical and economic implications, particularly regarding the choice between nutritional routes (enteral vs. parenteral) for patients in coronary intensive care, remain inadequately characterized. This study investigates the combined effects of nutritional route and nutritional risk on hospital costs and length of stay among patients in the coronary intensive care unit.

Methods: Data from 217 patients in the coronary intensive care unit (ICU) were analyzed. Linear regression and analysis of covariance were used to evaluate associations between nutritional approach and total hospitalization cost, adjusting for age, sex, C-reactive protein level, diabetes status, and length of stay.

Results: In the basic model, a longer length of stay was significantly associated with increased costs, whereas enteral nutrition (EN) was associated with lower costs. In alternative models, EN continued to limit cost escalation even during prolonged stays. The absence of nutritional risk was also associated with smaller increases in costs related to prolonged hospitalization. Patients who received EN, who had prolonged stays, and who demonstrated no nutritional risk experienced the greatest cost reductions.

Conclusion: The interaction between length of stay and nutritional approach is a major determinant of hospital costs for patients in the coronary ICU. Enteral nutrition appears to be associated with a more favorable cost trajectory during prolonged stays in the ICU. Routine nutritional risk assessment and appropriate selection of the nutritional support route may contribute to optimizing outcomes and controlling expenditures. Policies that encourage evidence-based use of EN and proactive nutritional risk management may yield meaningful clinical and economic benefits.

Keywords: Intensive care units, hospital costs, length of stay, risk assessment, risk management

Introduction

Intensive care units (ICUs) are specialized environments equipped with advanced resources and highly trained personnel to manage life-threatening conditions and monitor patients at risk of rapid clinical deterioration (1). Among the factors influencing outcomes in critically ill patients, malnutrition is a major determinant of morbidity, mortality, prolonged length of stay (LOS), and increased healthcare costs (2-4). Nutritional support is therefore a cornerstone of ICU care, aiming to provide adequate energy and nutrients through enteral nutrition (EN) or parenteral nutrition (PN) (5-7). A study has consistently demonstrated that optimized nutritional therapy improves clinical outcomes (8). International guidelines recommend early initiation of EN whenever clinically feasible because of its beneficial effects on gastrointestinal integrity, reduced risk of infection, shorter LOS, and lower healthcare costs

Corresponding Author: Asst. Prof., Cagla Ayer, Izmir Katip Celebi University Faculty of Health Sciences, Department of Nutrition And Dietetics, İzmir, Türkiye
E-mail: cagla.dalbay@gmail.com **ORCID:** orcid.org/0000-0001-6124-7339
Received: 23.01.2026 **Accepted:** 23.08.2026 **Publication Date:** 17.11.2026

Cite this article as: Ayer C, Demir H, Akar T, Mertoglu S. The impact of nutritional approach on length of stay and costs in coronary intensive care patients: a retrospective analysis. Med Bull Haseki. 2026;64(5):372-383

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

(9). In contrast, PN is primarily recommended when EN is contraindicated or fails to meet nutritional requirements, as PN is associated with a higher risk of metabolic and infectious complications and requires careful patient selection and close metabolic monitoring (10).

Despite these established benefits, evidence regarding the economic implications of nutritional strategies in coronary intensive care units (C-ICUs) remains limited. Patients admitted to C-ICUs have distinct hemodynamic and metabolic characteristics that may influence both nutritional management and resource utilization; yet previous studies have generally evaluated nutritional route and nutritional risk independently rather than considering their combined effects on hospital costs. We hypothesized that EN would be associated with lower hospital costs than PN; that higher nutritional risk would be associated with increased costs; and that the association between prolonged LOS and hospital costs would vary according to nutritional route, with a smaller increase in costs associated with longer hospitalization among patients receiving EN.

Accordingly, the primary objective of this study was to evaluate the association of the nutritional approach [defined by the route of nutritional support (EN vs. PN)] and nutritional risk with total hospital costs among patients admitted to the C-ICU. The secondary objective was to determine whether the relationship between LOS and hospital costs was modified by nutritional route and nutritional risk. Accordingly, the primary outcome was total hospital cost. The secondary outcomes were the interaction effects of nutritional route × LOS and nutritional risk × LOS on hospital costs and were estimated using analysis of covariance models adjusted for age, sex, C-reactive protein (CRP), diabetes status, and LOS. This study is expected to provide clinically relevant evidence supporting early nutritional risk assessment and the timely implementation of EN, thereby contributing to more cost-effective nutritional management and improved resource utilization in C-ICUs.

Materials and Methods

Compliance with Ethical Standards

Ethical approval was obtained from the Izmir Katip Celebi University Social Research Ethics Committee (approval number: 2025/24-07, date: 10.12.2025).

Study Design

This single-center, retrospective, observational, analytical study used routinely collected electronic health records of adult patients admitted to the coronary ICU. Patient data were obtained from the hospital records, with no intervention in clinical decision-making or alteration of patients' nutritional management. This design enabled the assessment of associations between nutritional support strategies, nutritional risk status, length of hospital stay, and total hospital costs within the context of routine clinical practice. This study was conducted from the perspective of the Turkish Social Security Institution (SSI), and all hospital costs reflected SSI reimbursement payments rather than actual hospital expenditures. The study population consisted of adult patients admitted to the ICU (secondary-level care) at Izmir Bakircay University Cigli Training and Research Hospital between January 1, 2020, and December 31, 2024. From this population, 217 adult patients with complete electronic records were included in the analysis. The overall study flow, including patient selection and analytical procedures, is presented in Figure 1. For context, Kalın et al. (11) examined 100 patients when investigating the relationship between nutritional risk and length of stay. As this was a retrospective study encompassing all eligible patients within the predefined study period, a priori sample size calculation was not conducted. Compared with this previous study, the inclusion of 217 patients in the present study provided an appropriate study population to examine the proposed relationships and to evaluate interaction effects between nutritional approach and length of stay. To further assess the adequacy of the available sample relative to the complexity of the final model, a sensitivity analysis was performed using G*Power 3.1.9.7 with eight model terms (seven main effects and one three-way interaction), an α level of 0.05, and 80% power. The minimum detectable effect size was determined to be $f^2=0.0365$.

This study employed a two-stage analytical framework designed to estimate baseline associations and evaluate how the association between LOS and cost varied by nutritional approach. The basic model used linear regression, while the alternative models applied analysis of covariance to estimate interaction effects. Because the outcome variable, total hospital cost, was continuous and heavily right-skewed, its distribution was examined using R (12), and substantial positive skewness was observed. Consistent with Alpar's recommendations for skewed cost data (13), a logarithmic transformation was applied to ensure compliance with multivariable model assumptions.

The key independent variables were route of nutritional support (EN vs. PN) and nutritional risk (present vs. absent), which together operationalized the nutritional approach. Both variables were modeled categorically. Covariates included age, sex, CRP, diabetes status, and LOS. These variables were selected based on their theoretical and empirical relevance to cost variation in ICU populations. Age was treated as a discrete numerical variable, LOS and CRP as continuous variables, and diabetes status as a categorical variable. The conceptual model of factors influencing costs in C-ICU patients is presented in Figure 2.

Figure 2 summarizes the conceptual framework linking demographic, hospital-related, chronic disease, infection-related, and nutritional factors to hospital costs. Each factor was assumed to exert a positive effect on cost because of higher resource utilization or greater clinical complexity. Specifically, longer LOS was expected to increase hospital costs because of extended treatment requirements; diabetes and elevated CRP were expected to increase hospital costs as indicators of greater disease severity; and, from a nutritional perspective, PN and nutritional risk [based on Nutritional Risk Screening 2002 (NRS-2002)] were expected to increase hospital costs.

Diabetes status was determined using ICD-10 code E11. The nutritional route was extracted from medication records: patients receiving enteral products were classified as EN, and those receiving parenteral products were classified as PN. No patient received EN and PN concurrently. Nutritional risk was assessed using NRS-2002, a validated and widely recommended screening tool for predicting adverse outcomes in ICU settings. Previous studies also confirmed NRS-2002 as a strong predictor of clinical trajectories in critically ill patients (3,14). Patients with NRS-2002 scores of ≥3 were classified as having nutritional risk, whereas those with scores <3 were classified as having no nutritional risk.

Statistical Analysis

Descriptive statistics summarized the sample using frequencies for categorical variables and means with standard deviations for continuous variables. Differences between groups were evaluated using the chi-square test for categorical variables and the independent samples t-test for continuous variables (15). Initial data exploration was performed using visual tools in R (12) to assess internal consistency. The findings of the initial data exploration, including the distribution of the outcome variable and

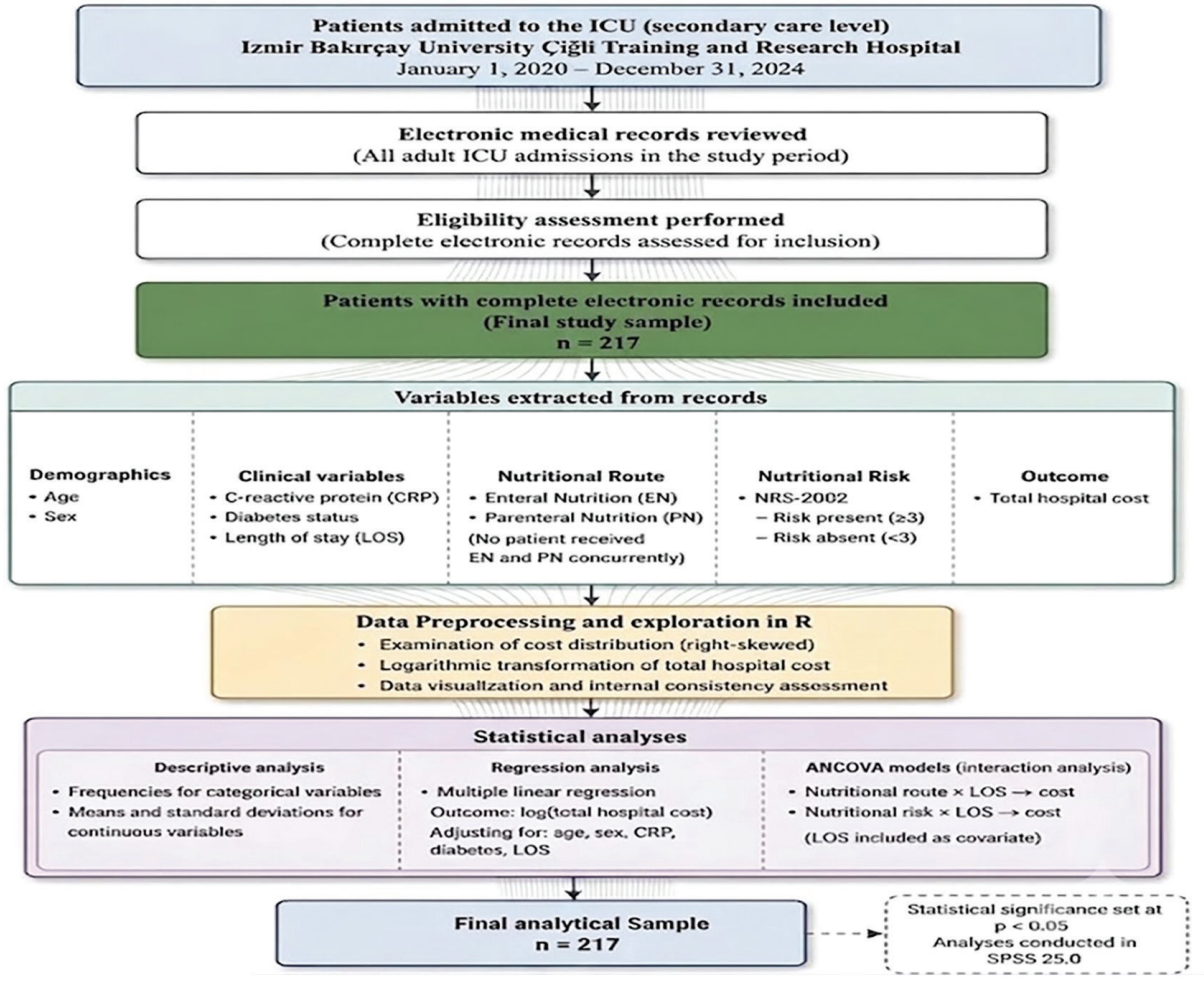


Figure 1. Study flow diagram and analytical framework
ICU: Intensive care unit; EN: Enteral nutrition; PN: parenteral nutrition; NRS-2002: Nutritional risk screening 2002; LOS: Length of stay; CRP: C-reactive protein

assessment of data characteristics, were presented in Figure 3.

Regression analysis estimated the association between nutritional approach and log-transformed hospital cost while adjusting for covariates (13,15,16). Covariance models were then used to evaluate whether the relationship between LOS and hospital costs varied by nutritional route and nutritional risk. The rationale for using analysis of covariance was its suitability for estimating moderated effects by incorporating LOS as a covariate alongside categorical predictors (17). The assumption of linearity between LOS and hospital cost was confirmed.

Data preprocessing and visualization were performed using R (12), whereas regression and covariance analyses were conducted using SPSS version 25.0 (IBM Corp., Armonk, NY, USA) (18). Statistical significance was set at $p<0.05$. All analyses were conducted under the assumption that hospital data were accurate and reliable. All cost figures were reported in US dollars (USD). Original cost data were recorded in Turkish lira (TRY) and converted using the 2024 annual average exchange rate published by the (Central Bank of the Republic of Türkiye: 1 USD=32.51 TRY), which was applied uniformly across the study period for consistency in reporting. This study utilized specialized AI tools to enhance its rigor and efficiency: Notebook LM (19) for literature synthesis, ChatGPT (20) for English translation, and Claude AI (21) for proofreading and linguistic refinement. The authors remain fully responsible for the final content and interpretations.

Results

Patient characteristics are presented in Table 1. Of the 217 C-ICU patients, the mean age was 75.17±13.25 years, 55.3% were male, and 11.5% had diabetes. PN was administered more frequently than EN (57.6% vs. 42.4%), and nutritional risk was present in 15.7% of patients. Mean LOS was 9.34±9.42 days and the mean total hospital cost was $544.34±$730.84. Patients receiving EN were significantly older and had longer hospital stays, higher total hospital costs, and higher daily costs than those receiving PN ($p=0.029$, $p<0.001$, $p=0.002$, and $p<0.001$, respectively).

Bivariate associations between key variables and cost are shown in Figure 3. Age, sex, diabetes status, and CRP were not significantly associated with cost (all $p>0.05$). LOS showed a significant positive correlation with cost ($r=0.504$; $p<0.001$). Costs also differed significantly by nutritional route, with higher unadjusted costs observed in the EN group ($t=4.798$; $p<0.01$), whereas no significant differences were observed by nutritional risk ($p=0.380$).

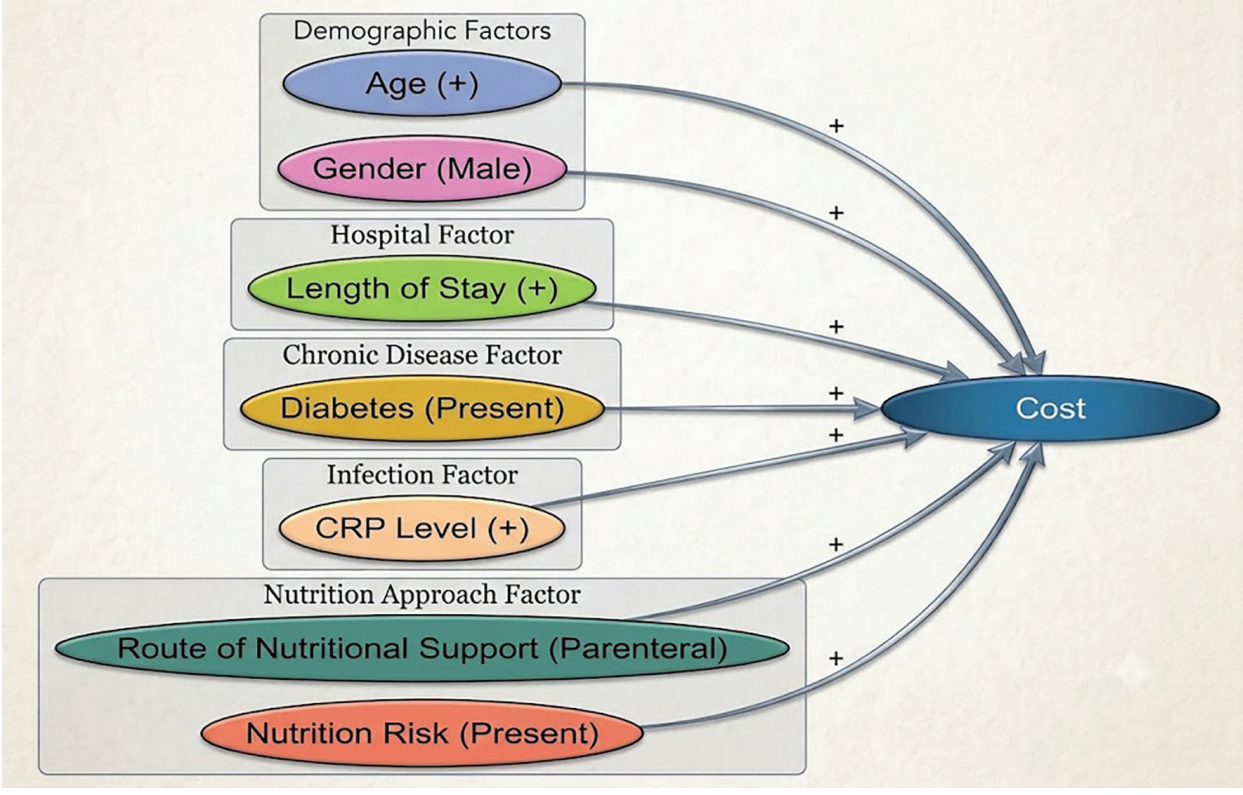


Figure 2. Conceptual model of factors affecting costs in C-ICU
C-ICU: Coronary intensive care unit

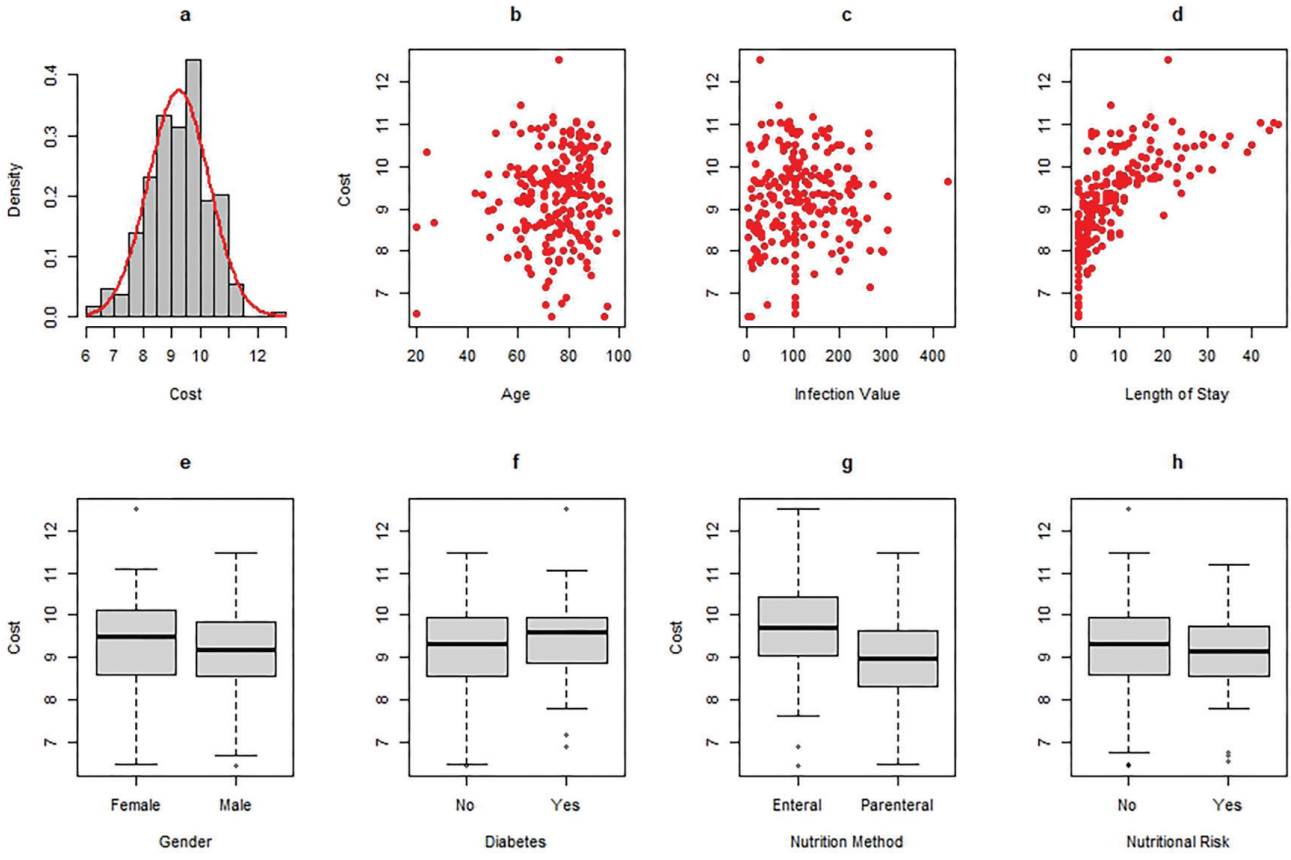


Figure 3. Findings on the relationship between clinical and demographic variables and costs in C-ICU patients. Associations between individual variables and hospitalization cost (log-transformed) were assessed using Pearson correlation for continuous variables and independent samples t-test for categorical variables. Panel (a) shows the distribution of the transformed cost variable; panels (b-h) show bivariate associations between individual variables and cost. Statistical significance was set at $p<0.05$

EN: Enteral nutrition, PN: Parenteral nutrition, CRP: C-reactive protein, LOS: Length of stay, C-ICU: Coronary intensive care unit

Table 2 presents the regression and covariance models that examine predictors of hospitalization costs. Across all models, LOS was the strongest and most consistent predictor of cost. In the basic regression model, after adjustment for age, sex, CRP, diabetes, and nutritional risk, LOS remained positively associated with cost (B=0.072; p=0.001). The nutritional route was significantly associated with cost in the opposite direction to that observed in the unadjusted analysis, and EN was associated with lower cost levels (B=-0.246; p=0.038). Age, sex, CRP, diabetes, and nutritional risk were not statistically significant predictors in this model, which explained 46.1% of the variance in cost (R^2=0.461).

The interaction models showed that the association between LOS and cost differed according to nutritional route and nutritional risk. The LOS × EN interaction was significant (B=-0.030; p=0.017), indicating that the marginal increase in cost associated with LOS was smaller among EN patients (Alternative Model 1). Similarly, the LOS × absence-of-nutritional-risk interaction was significant (B=-0.059; p=0.011), and CRP also reached significance in this model (B=0.019; p=0.045) (Alternative Model 2). The three-way interaction (LOS × EN × no nutritional risk) was significant in Alternative Model 3 (B=-0.071; p=0.015), identifying the smallest marginal cost increase among patients receiving EN without nutritional risk, and this model showed the best overall fit (R^2=0.488).

Discussion

This study investigated how nutritional route and nutritional risk shape hospitalization costs in C-ICU patients, with particular attention to their interaction with LOS. Across all models, LOS emerged as the dominant cost driver. However, the interaction analyses showed that the association between LOS and cost was not uniform but varied according to nutritional route and nutritional risk. This constitutes the central contribution of the study: that the nutritional route and nutritional risk modify the

Table 1. Descriptive findings on demographic and clinical variables							
Variables	**Total**		**EN**		**PN**		**p**
	n	**%**	**n**	**%**	**n**	**%**	
Route of nutritional support	217	100.0	92	42.4	125	57.6	-
Gender							
Female	97	44.7	48	52.2	49	39.2	0.072
Male	120	55.3	44	47.8	76	60.8	
Presence of diabetes							
Absent	192	88.5	84	91.3	108	86.4	0.290
Present	25	11.5	8	8.7	17	13.6	
Nutritional risk							
Absent	183	84.3	80	87.0	103	82.4	0.451
Present	34	15.7	12	13.0	22	17.6	
	$\bar{x}$ ± SD		**$\bar{x}$ ± SD**		**$\bar{x}$ ± SD**		
Age (year)	75.17±13.25		77.46±12.25		73.48±13.74		**0.029**
CRP (mg/dL)	115.45±74.62		112.16±63.58		117.87±81.96		0.564
Length of Stay (days)	9.34±9.42		12.92±10.84		6.70±7.18		**<0.001**
Cost (USD)	$544.34±730.84		$743.43±957.93		$397.81±453.81		**0.002**
Daily cost (USD/day)	$5.77±1.07		$6.16±1.00		$5.49±1.03		**<0.001**

Data are presented as mean ± SD for continuous variables and n (%) for categorical variables. EN: Enteral nutrition, PN: Parenteral nutrition, CRP: C-reactive protein, USD: United States dollar, SD: Standard deviation

marginal cost of prolonged hospitalization, revealing a pattern not consistently captured by main-effects-only analyses in prior research. This gap is particularly relevant for the C-ICU setting, where a recent review highlighted a lack of robust evidence regarding appropriate nutritional support despite the high prevalence of malnutrition and its adverse impact on prognosis in acute heart failure, acute coronary syndrome, cardiogenic shock, and post-cardiac arrest patients (22).

The interaction between EN and LOS indicated that EN was associated with a smaller incremental increase in costs as hospitalization lengthened. This pattern may reflect differences in clinical characteristics and resource utilization between patients receiving EN and PN. Factors such as clinical stabilization, infection-related complications, and the use of high-cost interventions may contribute to these differences; however, these factors were not directly assessed in this study, and residual and unmeasured confounding cannot be excluded. In contrast, PN is typically administered to clinically unstable patients and is associated with risks-such as catheter-related infections and metabolic disturbances-that may help explain the steeper cost trajectory observed with longer PN-associated admissions. Notably, current guidance for this specific population cautions that EN in hemodynamically unstable patients carries a low but recognized risk of intestinal ischemia, which likely contributes to clinicians' preference for PN in the most unstable patients (22)-a pattern directly consistent with the confounding by indication discussed below.

Similarly, the interaction between LOS and nutritional risk indicated that patients without nutritional risk had smaller marginal cost increases during extended hospitalization. This pattern is consistent with a large C-ICU cohort study, in which a higher risk of malnutrition assessed by the geriatric nutritional risk index was significantly linked to increased in-hospital mortality, as well as longer hospital stay and C-ICU stay, compared with patients at no nutritional risk (23). That study's use of subgroup interaction analyses-finding significant interactions between nutritional status and mortality across hypertension, arrhythmia, cardiac arrest, shock, and chronic kidney disease subgroups-lends further methodological support to an interaction-based rather than a main-effects-only approach to modeling nutritional risk in C-ICU populations, consistent with the analytic strategy adopted here. These findings should be situated within, but not simply mapped onto, the existing literature. A recent meta-analysis of 14 studies found that early EN was associated with modest reductions in ICU and hospital LOS (mean differences of 0.18 and 1.15 days, respectively) and fewer bloodstream infections compared with early PN (24). Our finding of a significant LOS × nutritional route interaction, rather than a comparably small main effect, suggests one possible explanation for the modest pooled estimate: if the cost and LOS advantage of EN are not constant but increase with duration of hospitalization, then studies that report only average effects across heterogeneous LOS distributions may understate its clinical and economic relevance for patients

Table 2. Findings related to the base and alternative models											
Variables	**SS**	**df**	**F**	**p**	**η^2**	**B**	**SE**	**t**	**p**	**95% CI**	**Model findings**
Basic model											
Age (year)	0.032	1	0.050	0.823	0.000	0.001	0.004	0.224	0.823	-0.008-0.010	R^2: 0.461 Adjusted R^2: 0.443 F: 25.575 (p<0.001)
CRP (mg/dL)	2.337	1	3.682	0.056	0.017	0.001	0.001	1.919	0.056	0.000-0.003	
Length of stay (days)	86.144	1	135.736	**<0.001**	0.394	0.072	0.006	11.651	**<0.001**	0.060-0.085	
Gender (Female)	0.872	1	1.374	0.243	0.007	0.136	0.116	1.172	0.243	-0.093-0.365	
Diabetes (absent)	0.039	1	0.061	0.805	0.000	0.043	0.173	0.247	0.805	-0.299-0.385	
Route of nutritional support (EN)	2.766	1	4.358	**0.038**	0.020	-0.246	0.118	-2.088	**0.038**	-0.478-(-0.014)	
Nutrition risk (absent)	0.060	1	0.094	0.760	0.000	0.047	0.153	0.306	0.760	-0.254-0.347	
Alternative model 1											
Age (year)	0.019	1	0.030	0.863	0.000	0.001	0.004	0.173	0.863	-0.008-0.009	R^2: 0.476 Adjusted R^2: 0.456 F: 23.613 (p<0.001)
CRP (mg/dL)	2.085	1	3.360	0.068	0.016	0.001	0.001	1.833	0.068	0.000-0.003	
Length of stay (days)	89.393	1	144.077	**<0.001**	0.409	0.092	0.010	9.134	**<0.001**	0.072-0.111	
Gender (female)	1.052	1	1.696	0.194	0.008	0.150	0.115	1.302	0.194	-0.077-0.376	
Diabetes (absent)	0.010	1	0.017	0.898	0.000	0.022	0.172	0.129	0.898	-0.317-0.361	
Route of nutritional support (EN)	6.313	1	10.176	**0.002**	0.047						
EN						0.502	0.161	3.118	**0.002**	0.185-0.819	
PN						0[a]					
Nutrition risk (absent)	0.076	1	0.123	0.726	0.001	0.053	0.151	0.351	0.726	-0.244-0.350	
Length of stay (days)× Route of nutritional support	3.588	1	5.782	**0.017**	0.027						
Length of stay (days)× Route of nutritional support (EN)						-0.030	0.013	-2.405	**0.017**	-0.005-0.027	
Length of stay (days)× Route of nutritional support (PN)						0[a]					
Alternative model 2											
Age (year)	0.003	1	0.004	0.949	0.000	0.000	0.004	-0.064	0.949	-0.009-0.008	R^2: 0.478 Adjusted R^2: 0.458 F: 23.805 (p<0.001)
CRP (mg/dL)	2.503	1	4.050	**0.045**	0.019	0.001	0.001	2.012	**0.045**	2.97E-5-0.003	
Length of stay (days)	43.729	1	70.753	**<0.001**	0.254	0.127	0.022	5.745	**<0.001**	0.084-0.171	
Gender (female)	0.715	1	1.156	0.283	0.006	0.123	0.115	1.075	0.283	-0.103-0.349	
Diabetes (absent)	0.059	1	0.095	0.758	0.000	0.053	0.171	0.309	0.758	-0.285-0.390	
Route of nutritional support (EN)	2.636	1	4.266	**0.040**	0.020	-0.240	0.116	-2.065	**0.040**	-0.469-(-0.011)	
Nutrition risk (absent)	1.603	1	2.593	0.109	0.012						
Absent						0.345	0.214	1.610	0.109	-0.077-0.767	
Present						0[a]					
Length of stay (days)× Nutrition risk	4.086	1	6.611	**0.011**	0.031						
Length of stay (days)× Nutrition risk (absent)						-0.059	0.023	-2.571	**0.011**	-0.103-(-0.031)	
Length of stay (days)× Nutrition risk (present)						0[a]					
Alternative model 3											
Age (year)	0.003	1	0.004	0.947	0.000	0.000	0.004	-0.066	0.947	-0.009-0.008	R^2: 0.488 Adjusted R^2: 0.463 F: 19.647 (p<0.001)
CRP (mg/dL)	2.247	1	3.673	0.057	0.018	0.001	0.001	1.917	0.057	-3.987-0.003	
Length of stay (days)	43.598	1	71.254	**<0.001**	0.257	0.131	0.028	4.749	**<0.001**	0.077-0.186	
Gender (female)	0.840	1	1.373	0.243	0.007	0.134	0.115	1.172	0.243	-0.092-0.360	

Table 2. Continued

Variables	SS	df	F	p	η^2	B	SE	t	p	95% CI	Model findings
Diabetes (absent)	0.024	1	0.039	0.844	0.000	0.034	0.171	0.197	0.844	-0.303-0.370	R^2: 0.488 Adjusted R^2: 0.463 F: 19.647 (p<0.001)
Route of nutritional support (EN)	4.872	1	7.962	**0.005**	0.037						
EN						0.463	0.164	2.822	**0.005**	0.140-0.787	
PN						0[a]					
Nutrition risk (absent)	1.055	1	1.724	0.191	0.008						
Absent						0.283	0.215	1.313	0.191	-0.142-0.708	
Present						0[a]					
Length of stay (days)× Route of nutritional support× Nutrition risk	6.597	3	3.594	**0.015**	0.050						
Length of stay (days)× Route of nutritional support (EN)× Nutrition risk (absent)						-0.071	0.029	-2.454	**0.015**	-0.129-(-0.014)	
Length of stay (days)× Route of nutritional support (EN)× Nutrition risk (present)						-0.017	0.033	-0.522	0.602	-0.082-0.047	
Length of stay (days)× Route of nutritional support (PN)× Nutrition risk (absent)						-0.046	0.029	-1.577	0.116	-0.103-0.011	
Length of stay (days)× Route of nutritional support (PN)× Nutrition risk (present)						0[a]					

[a]Reference category

Hospitalization costs were analyzed using a general linear model (univariate analysis of covariance, ANCOVA) with Type III sum of squares. For categorical variables, the reference category is indicated as "0[a]," and its coefficient is fixed at zero by definition. Statistical significance was set at p<0.05. The basic model was estimated using linear regression; alternative Models 1-3 were estimated using analysis of covariance (ANCOVA, Type III sum of squares) to evaluate interaction effects between length of stay and nutritional approach. The outcome variable (total hospital cost) was log-transformed prior to analysis due to significant positive skewness. For categorical variables, the reference category is indicated as "0[a]," with its coefficient fixed at zero by definition. The first p-value (adjacent to F) reflects the overall (Type III) test of the variable's effect; the second p-value (adjacent to t) reflects the significance of the individual regression coefficient (B). Statistical significance was set at p<0.05. SS: Sum of squares, df: Degrees of freedom, F: F Statistic, η^2: Partial eta squared, B: Unstandardized regression coefficient, SE: Standard error, t: t statistic, CI: Confidence interval, R^2: Coefficient of determination, CRP: C-reactive protein, EN: Enteral nutrition, PN: Parenteral nutrition, ×: denotes the interaction term between the variables

with prolonged stays. This interaction-based perspective may offer a more informative lens than reliance on main effects alone when evaluating nutritional routes in future cost analyses.

A systematic study demonstrated that withholding early PN reduced infections and accelerated recovery from organ failure, translating into shorter hospitalization and lower costs (25), and a cost-effectiveness analysis similarly concluded that early PN increased costs without improving outcomes relative to EN (26). These findings support the plausibility of a genuine, mechanism-driven cost advantage for EN. However, because these studies remove the influence of baseline illness severity on treatment assignment-precisely the factor our observational design could not fully account for-these trials cannot confirm that the magnitude of the association observed in our cohort reflects treatment effect alone. It is more likely that our findings reflect a combination of a modest treatment-related effect and a selection-driven component due to confounding by indication (discussed below); the relative contribution of each cannot be disentangled using the present data. Current guidelines recommending EN within the first 48 hours (27-29) further support the physiological rationale for early EN, though guideline endorsement reflects evidence from controlled settings that may not fully generalize to the heterogeneity of routine C-ICU practice.

The relatively low prevalence of nutritional risk in this cohort (15.7%) also warrants comment, as it is notably lower than rates reported in other ICU and cardiac surgery populations. While rates of NRS-2002-defined risk reported in general ICU and cardiac surgery populations are often considerably higher (approximately 40-46%) (30,31), a more directly comparable population-patients with acute coronary syndrome-shows a pattern more consistent with our findings. In a large retrospective study, NRS ≥3 was

present in only 6% of survivors, compared with 20% of those who died in hospital (32), indicating that nutritional risk by this definition is comparatively uncommon in these populations overall, though strongly associated with mortality when present. Given that survivors substantially outnumber non-survivors in most C-ICU cohorts, including ours, a population-wide prevalence in the range of 15-20% is plausible for this specific patient group and does not need to be solely attributable to incomplete NRS-2002 documentation, as previously suggested. This distinction matters clinically: even at low prevalence, nutritional risk identified by NRS-2002 appears to carry meaningful prognostic weight in these populations and, by extension, in C-ICU patients, reinforcing the value of systematic screening despite its low positive rate in these populations.

These associations must be interpreted with caution, given the observational and non-randomized nature of assignment to nutritional routes. In routine clinical practice, the decision to initiate EN versus PN is guided by the patient's clinical status rather than made independently of it: patients who are hemodynamically unstable, have gastrointestinal contraindications, or are otherwise more severely ill are more likely to receive PN, whereas clinically stable patients are more likely to receive EN. This selection pattern aligns with current clinical guidance for this population, which recommends EN over PN as a general principle while explicitly acknowledging the elevated-if low absolute-risk of intestinal ischemia with EN in hemodynamically unstable patients, alongside preference for hypocaloric intake in the first week of ICU admission and vigilance for refeeding syndrome in malnourished patients (22). These are precisely the clinical considerations that would have guided route selection in our cohort but could not be captured by the covariates available in this study. Notably, the unadjusted analysis showed higher costs in the EN group, whereas the adjusted models showed the opposite: EN was associated with lower costs. This reversal of direction after adjustment is itself consistent with confounding by indication: it suggests that unadjusted comparisons may have been dominated by baseline differences in illness severity between groups and that even the adjusted estimate should be interpreted cautiously, since the covariates available (age, sex, CRP, diabetes, and LOS) may not fully capture the clinical basis on which the nutritional route was chosen. In particular, the models did not include several variables known to jointly influence both nutritional route selection and hospitalization cost, including validated severity-of-illness scores (APACHE II/IV, SOFA, SAPS II), indicators of organ support and hemodynamic instability (mechanical ventilation, vasopressor use), cardiac functional status (e.g., ejection fraction), primary admission diagnosis (e.g., myocardial infarction, heart failure, arrhythmia), and overall comorbidity burden (Charlson Comorbidity Index). The absence of these variables substantially raises the risk of residual confounding, as patients selected for PN are likely to differ systematically from those selected for EN on these dimensions. Consequently, the observed associations between the nutritional route and cost, including interaction effects with LOS, may partly reflect underlying differences in patient acuity and diagnoses that were not fully captured in the available data and should therefore not be interpreted as causal effects of the nutritional approach. While the consistency of this pattern across multiple models and its partial alignment with mechanistic evidence from randomized trials lend some support to a genuine association, the possibility that residual confounding accounts for part or all of the observed effect cannot be excluded. Future studies incorporating these severity and diagnostic indicators, along with propensity score-based methods, would help clarify the extent to which this association reflects a treatment-related effect versus underlying patient selection.

The findings of this study should be interpreted within the context of the SSI perspective. Unlike studies conducted from hospital, societal, or healthcare system perspectives, the cost estimates in the present study were based on standardized SSI reimbursement payments rather than on actual hospital expenditures or resource consumption. Therefore, differences in absolute cost values compared with those reported in the international literature are expected, primarily reflecting variations in reimbursement policies, healthcare financing systems, costing methodologies, and country-specific pricing structures. Nevertheless, the observed associations between nutritional approach, nutritional risk, LOS, and hospital costs remain clinically and economically relevant, as they provide insight into the relationships among these factors within the Turkish reimbursement system and may inform future reimbursement policies and resource allocation decisions.

Together, these findings suggest that the cost implications of nutritional support cannot be understood solely through main effects. Instead, the nutritional route and nutritional risk appear to interact with LOS in clinically relevant ways, shaping observed patterns of resource utilization. For clinicians, these associations may support the value of systematic nutritional risk assessment and consideration of EN in eligible patients, when clinically appropriate, in line with existing guidelines, while recognizing that individual clinical judgment regarding EN feasibility must remain paramount. For managers and policymakers, strategies that support EN availability, promote guideline adherence, and integrate nutritional screening into routine ICU protocols may be worth further evaluation as potential approaches to managing ICU expenditures, pending confirmation in studies better able

to account for illness severity. While variables such as age, sex, CRP, and diabetes showed limited direct associations, the interplay between nutritional approach and LOS emerged as the most clinically relevant correlate of cost in this cohort. These findings are hypothesis-generating and suggest that EN-particularly in patients without nutritional risk-warrants further prospective investigation, ideally incorporating adjustment for severity or randomization, as a potential strategy for mitigating the economic burden of care in the coronary ICU.

Study Limitations

This study has several limitations that should be considered when interpreting the findings. Although the dataset included 217 C-ICU patients, incomplete information for some variables, particularly the absence of recorded NRS-2002 assessments, reduced analytical precision. This limitation reflects broader challenges in routine malnutrition screening and the integration of nutritional assessment tools into electronic medical record systems and calls for improved clinical documentation practices. Second, the categorization of patients into EN and PN groups restricted the evaluation of combined nutritional approaches. Because very few patients received simultaneous EN and PN, the study could not assess how hybrid nutrition strategies might influence LOS or costs. As combined approaches are increasingly used in modern ICU nutrition therapy, future research with more granular data is warranted. Third, the analysis relied on hospital-reported data, whose accuracy was assumed. Although this approach is standard in retrospective studies, it introduces potential risks related to documentation inconsistencies or coding variability. Fourth, perhaps most importantly, the allocation of patients to EN versus PN was not randomized but rather reflected clinical decision-making, introducing the possibility of confounding by indication. The multivariable models adjusted only for age, sex, CRP, diabetes, and LOS and did not include several clinically important variables known to influence both nutritional route selection and hospitalization cost. These include validated severity-of-illness scores (APACHE II/IV, SOFA, SAPS II), indicators of organ support and hemodynamic instability (mechanical ventilation, vasopressor use), cardiac functional status (ejection fraction or other measures of cardiac function), primary admission diagnosis (e.g., myocardial infarction, heart failure, arrhythmia), and overall comorbidity burden (Charlson Comorbidity Index). The absence of these variables substantially raises the risk of residual confounding, as patients selected for PN are likely to differ systematically from those selected for EN with respect to these dimensions, for example, greater illness severity, hemodynamic instability, or gastrointestinal contraindications to EN. Consequently, the observed associations between the nutritional route and cost, including the interaction effects with LOS, may partly reflect underlying differences in patient acuity and diagnoses that were not fully captured in the available data and therefore should not be interpreted as direct causal effects of the nutritional approach. Future studies should incorporate these severity and diagnostic indicators, along with propensity-score-based methods, to more rigorously isolate the contribution of the nutritional route from that of the underlying clinical status. Fifth, three separate covariance models were constructed to evaluate interaction effects involving LOS, nutritional route, and nutritional risk. No correction for multiple comparisons (e.g., Bonferroni or false discovery rate adjustment) was applied across these models, increasing the risk of Type I error. Although the models were built sequentially to address distinct, conceptually related questions rather than as independent hypothesis tests, applying a conservative correction would render the LOS × EN interaction in Alternative model 1 ($p=0.017$) marginally non-significant, whereas the interactions in Alternative Models 2 and 3 would remain significant. Future studies should consider formal multiplicity adjustment or preregistered analysis plans to more rigorously control for this risk. Sixth, because this was a retrospective study encompassing all eligible patients during the predefined study period, an a priori sample size calculation was not performed. A sensitivity analysis indicated that relatively small effects ($f^2 \geq 0.0365$) could be detected with 80% power, suggesting that the available sample was adequate for evaluating the proposed relationships; nonetheless, the absence of prospective power planning should be acknowledged as a methodological limitation.

Taken together, these limitations should be considered when interpreting the findings but do not diminish the study's overall contribution. Instead, they highlight important directions for improving data quality and for guiding future research on nutritional strategies in C-ICU settings. Despite these limitations, the present study has several important strengths. It is one of the few studies to simultaneously evaluate the nutritional route, nutritional risk, and their interaction with LOS with respect to hospital costs among coronary ICU patients. Furthermore, it used detailed patient-level electronic health records collected over five years and applied multivariable statistical models to adjust for key clinical confounders, thereby providing a comprehensive assessment of the economic implications of nutritional management in routine clinical practice.

Conclusion

The present study demonstrated that EN was associated with lower hospital costs during prolonged C-ICU stays, particularly among patients who underwent systematic

nutritional risk assessment. These findings suggest that incorporating nutritional risk screening into nutritional management and considering EN as a nutritional support strategy may contribute to more efficient resource utilization; however, given the observational design and the potential for residual confounding, a causal interpretation of these findings is not warranted. Prospective multicenter studies are needed to evaluate standardized EN protocols and their associated clinical and economic outcomes.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Izmir Katip Celebi University Social Research Ethics Committee (approval number: 2025/24-07, date: 10.12.2025).

Informed Consent: This study was designed as a retrospective observational study using routinely collected electronic health records. Patient data were obtained from hospital records, with no intervention in clinical decision-making or alteration of patients' nutritional management. This design enabled the assessment of associations between nutritional support strategies, nutritional risk status, length of hospital stay, and total hospital costs within the context of routine clinical practice. Informed consent was not obtained from the patients due to the retrospective nature of the study.

Footnotes

Authorship Contributions

Concept: C.A., H.D., Design: C.A., H.D., Data Collection or Processing: C.A., H.D., T.A., S.M., Analysis or Interpretation: C.A., H.D., T.A., S.M., Literature Search: C.A., H.D., Writing: C.A., H.D.

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

Declaration Regarding the Use of AI and AI-Assisted Technologies: Artificial intelligence-assisted tools were used for language editing, manuscript formatting, table preparation, and literature checking. All outputs were reviewed and verified by the authors. The authors take full responsibility for the content of the manuscript, the accuracy of the data, and the integrity of the work.

References

1. Ediboğlu Ö, Moçin ÖY, Özyılmaz E, et al. Current statement of intensive care units in Turkey: data obtained from 67 centers. Turk Thorac J. 2018;19:209-15.
2. Lew CCH, Yandell R, Fraser RJL, Chua AP, Chong MFF, Miller M. Association between malnutrition and clinical outcomes in the intensive care unit: a systematic review. JPEN J Parenter Enteral Nutr. 2017;41:744-58.
3. Moghaddam OM, Hosseinzadeh Emam M, Irandoost P, et al. Relation between nutritional status on clinical outcomes of critically ill patients: emphasizing nutritional screening tools in a prospective cohort investigation. BMC Nutrition. 2024;10:69.
4. Xu J, Luo D, Chi R, et al. Comparison between the NUTRIC score and modified NUTRIC score to predict hospital mortality in patients undergoing cardiac surgery: a retrospective study. Nutr Clin Pract. 2026;41:166-76.
5. Brisard L, Le Gouge A, Lascarrou JB, et al. Impact of early enteral versus parenteral nutrition on mortality in patients requiring mechanical ventilation and catecholamines: study protocol for a randomized controlled trial (NUTRIREA-2).Trials. 2014;15:507.
6. Ido T, Hayashi T, Yamada T, Nakamoto K, Sotomi Y, Higuchi Y. Implementation of an enteral nutrition protocol in the cardiac care unit- a retrospective study of critically Ill cardiovascular patients. Circ Rep. 2025;8:127-35.
7. Yeh DD. Fueling recovery: evidence-based ICU nutrition and immunonutrition strategies in 2026. Trauma Surgery Acute Care Open. 2026;11:e002284.
8. Reignier J, Gaillard-Le Roux B, Dequin PF, et al. Expert consensus-based clinical practice guidelines for nutritional support in the intensive care unit: the French Intensive Care Society (SRLF) and the French-Speaking Group of Pediatric Emergency Physicians and Intensivists (GFRUP). Ann Intensive Care. 2025;15:99.
9. Alsharif DJ, Alsharif FJ, Aljuraiban GS, Abulmeaty MMA. Effect of supplemental parenteral nutrition versus enteral nutrition alone on clinical outcomes in critically ill adult patients: a systematic review and meta-analysis of randomized controlled trials. Nutrients. 2020;12:2968.
10. Blaser AR, Thibault R, Bachmann KF. Parenteral nutrition in the ICU. Clinical Nutrition Open Science. 2026;66:100635.
11. Kalın BS, Altun K, Sert Aİ, Kurt Çelebi N, Aktiz Bıçak E, Savuşma B. Investigation of the effect of the modified nutrition risk in critically ill score on the length of stay in the intensive care unit. Clin Sci Nutr. 2023;5:58-62.
12. R Core Team. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing; 2025.
13. Alpar R. Uygulamalı çok değişkenli istatistiksel yöntemler. 4. Baskı. Ankara: Detay Yayıncılık; 2013.
14. Popiołek-Kalisz J, Chromiński T, Szczasny M, Błaszczyk P. Nutritional status predicts the length of stay and mortality in patients undergoing electrotherapy procedures. Nutrients. 2024;16:843.
15. Alpar R. Spor, sağlık ve eğitim bilimlerinden örneklerle uygulamalı istatistik ve geçerlik-güvenirlik. 4. Baskı. Ankara: Detay Yayıncılık; 2016.
16. Gürbüz S, Şahin F. Sosyal bilimlerde araştırma yöntemleri. Ankara: Seçkin Yayıncılık; 2014.
17. Crawley MJ. The R book. England: Wiley Publications; 2009.
18. IBM Corp. IBM SPSS Statistics for Windows (Version 25.0) [Computer software]. IBM Corp; 2017. Available from: https://www.ibm.com/products/spss-statistics

19. OpenAI. Notebook LM [large language model]. OpenAI; 2025. Available from: https://notebooklm.google.com
20. OpenAI. ChatGPT (version 4) [AI model]. OpenAI; 2025. Available from: https://chat.openai.com
21. Anthropic. Claude AI (version 1.0) [AI language model]. Anthropic; 2025. Available from: https://claude.ai
22. Frederiks P, Peetermans M, Wilmer A. Nutritional support in the cardiac intensive care unit. Eur Heart J Acute Cardiovasc Care. 2024;13:373-9.
23. Li Y, Wang Z, Sun T, Zhang B, Liang X. Geriatric nutritional risk index was associated with in-hospital mortality among cardiac intensive care unit patients. Front Nutr. 2023;10:1218738.
24. Zhang G, Zhang K, Cui W, Hong Y, Zhang Z. The effect of enteral versus parenteral nutrition for critically ill patients: a systematic review and meta-analysis. JJ Clin Anesth. 2018;51:62-92.
25. Shi J, Wei L, Huang R, Liao L. Effect of combined parenteral and enteral nutrition versus enteral nutrition alone for critically ill patients: a systematic review and meta-analysis. Medicine. 2018;97:e11874.
26. Pradelli L, Graf S, Pichard C, Berger MM. Supplemental parenteral nutrition in intensive care patients: a cost-saving strategy. Clin Nutr. 2018;37:573-9.
27. Singer P, Blaser AR, Berger MM, et al. ESPEN practical and partially revised guideline: clinical nutrition in the intensive care unit. Clin Nutr. 2023;42:1671-89.
28. Reintam Blaser A, Starkopf J, Alhazzani W, et al. Early enteral nutrition in critically ill patients: ESICM clinical practice guidelines. Intensive Care Med. 2017;43:380-98.
29. Doğanay M, Akçay K, Çil T, et al. KEPAN enteral beslenme rehberi. Clin Sci Nutr. 2023;5:1-29.
30. Klek S, Szybinski P, Szczepanek K. Perioperative immunonutrition in surgical cancer patients: a summary of a decade of research. World J Surg. 2014;38:803-12.
31. Yang WB, Zhang W, Zhang JF, Yang TY, Mo ZF. Effect of supplementary parenteral nutrition on glucose and lipid metabolisms, risk of infection, and prognosis in critically ill patients with low body weight. Dig Med Res. 2019;2:10.
32. Czapla M, Karniej P, Juárez-Vela R, Łokieć K. The association between nutritional status and in-hospital mortality among patients with acute coronary syndrome-a result of the Retrospective Nutritional Status Heart Study (NSHS). Nutrients. 2020;12:3091.

Original Article

DOI: 10.4274/haseki.galenos.2026.25238
Med Bull Haseki 2026;64(5):384-390

Association of SGLT-2 Inhibitor Use with Wound Healing and Amputation Risk in Patients with Diabetic Foot Ulcers: A Prospective Observational Study

Zeynep Gizem Totik[1], Sevinc Puren Yucel Karakaya[2], Tayyibe Saler[3]

[1]*Adana Yuregir State Hospital, Ministry of Health, Clinic of Internal Medicine, Adana, Türkiye*
[2]*Cukurova University Faculty of Medicine, Department of Bioistatistics, Adana, Türkiye*
[3]*University of Health Sciences Türkiye, Adana City Training and Research Hospital, Clinic of Internal Medicine, Adana, Türkiye*

Abstract

Aim: The extant literature contains conflicting findings with regard to the safety of Sodium-glucose cotransporter-2 inhibitors (SGLT-2 inhibitors) in patients with diabetic foot ulcers (DFUs). The present study aims to evaluate the association between the use of SGLT-2 inhibitors and wound healing outcomes and the risk of amputation in patients diagnosed with type 2 diabetes mellitus (DM) and Wagner stage 1-2 DFU.

Methods: This prospective observational study was conducted between March 1, 2023, and August 31, 2024. The study population comprised 208 patients with type 2 DM and Wagner stage 1-2 DFU, all of whom completed six months of follow-up. Patients were grouped according to treatment: metformin plus insulin (n=102) or metformin plus insulin and an SGLT-2 inhibitor (n=106). Clinical, biochemical, wound-healing, and amputation outcomes were compared between groups. Subgroup analyses were performed for users of dapagliflozin and empagliflozin. Associations of SGLT-2 inhibitor use with wound healing and amputation were assessed using ordinal and binary logistic regression, respectively, with multivariable adjustment for relevant baseline differences.

Results: No statistically significant differences in wound healing and amputation rates were observed between patients with and without SGLT-2 inhibitor use. In the SGLT-2 inhibitor group, levels of serum albumin were elevated, whereas urinary albumin excretion was reduced ($p=0.002$ and $p<0.001$, respectively). No significant differences were observed between patients treated with dapagliflozin and empagliflozin in terms of wound healing or amputation.

Conclusion: No clear association has been observed between the use of SGLT-2 inhibitors and impaired wound healing or amputation in patients with DFUs. Further studies with larger sample sizes are needed to clarify these findings.

Keywords: Diabetes mellitus, type 2, foot ulcer, sodium-glucose transporter 2 inhibitors, wound healing

Introduction

Diabetes mellitus (DM) is a metabolic disorder characterized by hyperglycemia and multisystemic effects; the most common subtype is type 2 DM. DM has been shown to be associated with a wide range of macrovascular and microvascular complications (1). Diabetic foot ulcers (DFUs) represent a prevalent and grave complication associated with type 2 DM. As demonstrated in the extant literature, there is a strong correlation between DFUs and an elevated risk of infection, lower limb amputation, and death (2,3). The primary objective in the management of DM and its complications must be to select medications that exhibit a minimal side effect profile and are efficacious in addressing the pathophysiology underlying the condition. Sodium-glucose cotransporter-2 inhibitors (SGLT-2 inhibitors) constitute a class of oral antidiabetic agents recommended for suitable patients because of their beneficial cardiovascular, renal, and

Corresponding Author: Zeynep Gizem Totik, MD, Adana Yuregir State Hospital, Ministry of Health, Clinic of Internal Medicine, Adana, Türkiye
E-mail: zeynepgizemtotik@gmail.com **ORCID:** orcid.org/0000-0002-7742-7075
Received: 14.05.2026 **Accepted:** 23.08.2026 **Publication Date:** 17.11.2026

Cite this article as: Totik ZG, Yucel Karakaya SP, Saler T. Association of SGLT-2 inhibitor use with wound healing and amputation risk in patients with diabetic foot ulcers: a prospective observational study. Med Bull Haseki. 2026;64(5):384-390

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

metabolic effects (4). However, concerns have been raised regarding a potential increased risk of lower-extremity amputation, particularly following the CANVAS program, which reported a higher amputation rate with canagliflozin (5). In contrast, subsequent studies with dapagliflozin and empagliflozin have not consistently demonstrated a similar risk, leading to ongoing debate in the literature. In addition, the potential effects of SGLT-2 inhibitors on wound healing remain unclear (6). The present study examined the impact of the use of SGLT-2 inhibitors (i.e., empagliflozin and dapagliflozin) on wound healing and amputation in patients with type 2 DM and Wagner stage 1-2 DFUs. We hypothesized that wound-healing and amputation outcomes would differ based on SGLT-2 inhibitor use among patients with Wagner grade 1-2 DFUs. This study aims to contribute to the existing literature by comparing SGLT-2 inhibitor use among patients with DFUs.

Materials and Methods

Compliance with Ethical Standards

The study was approved by the Clinical Research Ethics Committee of University of Health Sciences Türkiye, Adana City Training and Research Hospital (decision no: 2372, date: 12.01.2023). The study was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki, and written informed consent was obtained from all participants.

Study Design

This was a prospective observational study conducted at the Clinic of Internal Medicine Adana City Training and Research Hospital between March 1, 2023, and August 31, 2024, with a 6-month follow-up period. The initial phase of the study involved 222 patients. During follow-up, five patients died and nine were lost to follow-up. Therefore, the final analysis included 208 patients (Figure 1). The study consisted of 102 patients receiving metformin plus insulin therapy and 106 patients treated with metformin, insulin, and an SGLT-2 inhibitor.

Data Collection and Clinical Assessment

Baseline demographic characteristics, duration of diabetes, body mass index, glycated hemoglobin (HbA1c) levels, complete blood count, and biochemical parameters were recorded. Hyperbaric oxygen therapy status was also documented.

The primary outcome of the study was wound healing, which was evaluated based on changes in Wagner stage and classified into three categories: regression (A) (improvement to a lower stage), no change (remaining the same stage) (B), and progression (C) (advancement to a higher stage). The

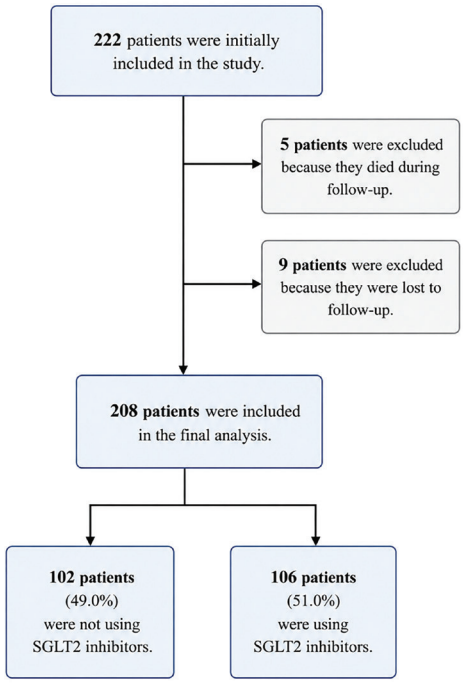


Figure 1. Flowchart of patient selection
SGLT-2 inhibitors: Sodium-glucose cotransporter-2 inhibitors

secondary outcome variable was defined as the occurrence of lower-extremity amputation during the follow-up period. All patients received education on wound care and nutrition in accordance with institutional protocols and were referred to relevant physicians for antibiotic therapy.

Study Groups

The division of patients into two groups was based on treatment regimen:

(1) Metformin plus insulin, and

(2) Metformin plus insulin plus an SGLT-2 inhibitor (dapagliflozin or empagliflozin).

Inclusion and Exclusion Criteria

Inclusion criteria were as follows: age ≥18 years, diagnosis of type 2 DM, presence of Wagner stage 1-2 DFUs, treatment with metformin plus insulin with or without an SGLT-2 inhibitor, and absence of significant arterial or venous stenosis (<50%). Patients who tolerated the treatment and had not discontinued it due to other adverse effects were included in the study; therefore, patients in the SGLT-2 inhibitor group were required to have received this treatment for at least six months prior to study inclusion.

Exclusion criteria included age <18 years, pregnancy, malignancy, rheumatologic disease, end-stage renal disease, prior amputation, and psychiatric disorders impairing self-care. In addition, patients with advanced

DFUs (Wagner ≥3), severe arterial or venous stenosis (≥50%), or conditions such as elephantiasis were excluded. Arterial and venous stenoses were assessed using color Doppler ultrasonography.

Statistical Analysis

Statistical analyses were conducted using IBM SPSS Statistics for Windows, version 20.0 (IBM Corp., Armonk, NY, USA). Descriptive data were summarized as frequency and percentage for categorical variables and as mean ± standard deviation or median (minimum-maximum) for continuous variables, depending on data distribution. Patients were compared by SGLT-2 inhibitor use (yes/no) and subtype (dapagliflozin vs. empagliflozin). Associations between categorical variables were evaluated using the chi-square test. For wound healing, assessed as an ordinal outcome (improvement, no change, and progression), a common odds ratio (OR) and its 95% confidence interval (CI) were estimated using ordinal logistic regression. For amputation, an OR with a 95% CI was calculated using binary logistic regression. Additionally, multivariable-adjusted analyses were performed for both outcomes to account for significant baseline differences in serum albumin, sodium, and the spot urine albumin-to-creatinine ratio. ORs with 95% CIs were reported. Normality of continuous variables was assessed with the Kolmogorov-Smirnov test. Continuous variables with a normal distribution were analyzed using the independent-samples t-test, whereas non-normally distributed variables were compared using the Mann-Whitney U test. Patients who discontinued SGLT-2 inhibitor treatment during follow-up remained in the analysis according to their treatment group at study entry. A two-sided p-value <0.05 was considered statistically significant.

Results

The initial phase of the study involved 222 patients. During follow-up, five patients died and nine were lost to follow-up. Therefore, the final analysis was performed on 208 patients. The study consisted of 102 patients receiving metformin plus insulin therapy and 106 patients treated with metformin, insulin, and an SGLT-2 inhibitor (Figure 1). Table 1 presents the baseline demographic, clinical, and laboratory characteristics of the study population in accordance with treatment groups. Significant differences were observed in spot urine albumin-to-creatinine ratio, serum albumin, and sodium levels (p=0.002, p=0.001, and p=0.017, respectively). Patients receiving SGLT-2 inhibitors had significantly lower spot albuminuria and higher serum albumin levels.

At the 6-month follow-up, no statistically significant between-group differences were observed for wound healing (common OR=0.64, 95% CI: 0.39-1.07, p=0.090) for lower-extremity amputation (OR=0.67, 95% CI: 0.35-1.25, p=0.264) (Table 2, Figure 2). After adjustment for serum albumin, sodium, and spot urine albumin-to-creatinine ratio, SGLT-2 inhibitor use remained not significantly associated with wound healing (aOR=0.77, 95% CI: 0.45-1.30, p=0.324) or with amputation (aOR=0.81, 95% CI: 0.42-1.55, p=0.522).

Table 3 presents a comparison of baseline demographic, clinical, and laboratory characteristics among patients receiving SGLT-2 inhibitors, stratified by inhibitor type [dapagliflozin (n=61) vs. empagliflozin (n=45)].

Table 4 shows that there were no major differences in wound healing or in the number of amputations among patients taking either dapagliflozin or empagliflozin, both SGLT-2 inhibitors (Figure 3).

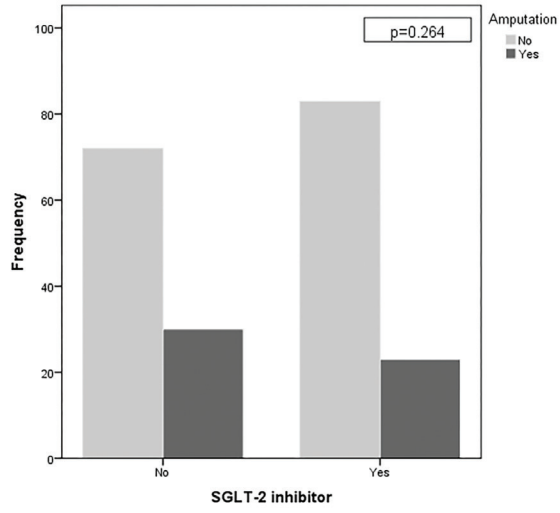


Figure 2. Amputation rates according to SGLT-2 inhibitor use
SGLT-2 inhibitors: Sodium-glucose cotransporter-2 inhibitors

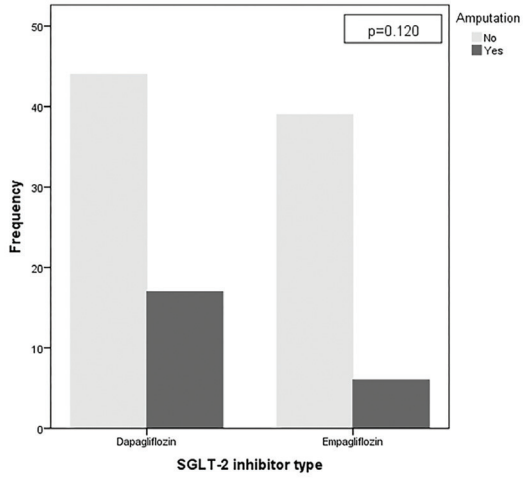


Figure 3. Amputation rates according to SGLT-2 inhibitor type
SGLT-2 inhibitors: Sodium-glucose cotransporter-2 inhibitors

Table 1. Demographic, clinical, and laboratory characteristics across SGLT-2 inhibitor groups			
	SGLT-2 inhibitors		p
	No (n=102)	Yes (n=106)	
Age, years	61.5 (39.0-82.0)	61.0 (36.0-76.0)	0.384
Sex, n (%)			0.220
Male Female	64 (62.7) 38 (37.3)	75 (70.8) 31 (29.2)	
Wagner grade, n (%)			0.299
Grade 1 Grade 2	38 (37.3) 64 (62.7)	47 (44.3) 59 (55.7)	
BMI (kg/m^2)	28.9 (22.6-40.3)	27.1 (22.7-37.8)	0.314
ALT (U/L)	14.0 (3.0-64.0)	16.0 (4.0-61.0)	0.129
AST (U/L)	16.0 (8.0-68.0)	16.0 (8.0-57.0)	0.732
Hemoglobin (g/dL)	11.9±1.8	12.0±1.9	0.597
WBC (×10^3/μL)	9.6 (3.6-28.9)	9.0 (3.9-33.0)	0.086
Neutrophil count (×10^3/μL)	6.6 (2.0-25.8)	6.1 (2.2-30.0)	0.098
Lymphocyte count (×10^3/μL)	1.9 (0.5-6.0)	1.9 (0.4-4.4)	0.658
Monocyte count (×10^3/μL)	0.65 (0.1-2.2)	0.65 (0.1-1.8)	0.827
Platelet count (×10^3/μL)	308.0 (145.0-663.0)	291.0 (140.0-690.0)	0.523
Serum albumin (g/dL)	3.7 (1.9-4.6)	4.0 (1.6-4.7)	0.001
eGFR (mL/min/1.73 m^2)	90.0 (40.0-128.0)	90.0 (38.0-132.0)	0.514
eGFR			0.777
≥90 mL/min/1.73 m^2 <90 mL/min/1.73 m^2	53 (52.0) 49 (48.0)	53 (50.0) 53 (50.0)	
Urea (mg/dL)	37.0 (13.0-94.0)	38.0 (16.0-138.0)	0.575
Creatinine (mg/dL)	0.9 (0.4-1.7)	0.9 (0.4-2.2)	0.999
C-reactive protein (mg/L)	27.5 (3.0-327.0)	30.0 (1.0-270.0)	0.893
Sodium (mmol/L)	137.5 (124.0-147.0)	139.0 (127.0-141.0)	0.017
Potassium (mmol/L)	4.5±0.4	4.5±0.5	0.797
HbA1c-baseline	9.6 (5.1-16.4)	9.6 (6.3-18.2)	0.637
HbA1c-6th month	9.0 (5.2-15.9)	8.9 (5.8-17.0)	0.712
Duration of diabetes, years	9.5 (2.0-28.0)	8.0 (1.0-20.0)	0.142
Spot urine albumin-to-creatinine ratio (mg/g)	150.0 (0.0-5579.0)	75.0 (2.0-4239.0)	0.002
Hyperbaric oxygen therapy, n (%)			0.746
No Yes	91 (89.2) 11 (10.8)	92 (86.8) 14 (13.2)	

Unless otherwise indicated, variables are expressed as mean ± standard deviation or median (minimum-maximum). Categorical variables were compared according to SGLT-2 inhibitor use using the chi-square test. Continuous variables are presented as mean ± standard deviation and were compared using the independent-samples t-test, or as median (min-max) and were compared using the Mann-Whitney U test, as appropriate

BMI: Body mass index, ALT: Alanine aminotransferase, AST: Aspartate aminotransferase, WBC: White blood cell count, eGFR: Estimated glomerular filtration rate, HbA1c: Glycated hemoglobin, SGLT-2 inhibitors: Sodium-glucose cotransporter-2 inhibitors

Discussion

According to the study, neither empagliflozin nor dapagliflozin was clearly associated with impaired wound healing or increased amputation rates in patients with Wagner stage 1-2 ulcers. A similar result was obtained in a single-center retrospective study by Dumortier et al. (7). In Dumortier et al.'s (7) study, patients with DFUs receiving SGLT-2 inhibitors were compared with patients not receiving SGLT-2 inhibitors. Over a six-month period, there was no clear association between wound healing rates in groups receiving SGLT-2 inhibitors and those not receiving SGLT-2 inhibitors. Furthermore, during the one-year follow-up period, the use of SGLT-2 inhibitors was not clearly associated with an increased risk of amputation. Similar to our study, this study included patients who received empagliflozin or dapagliflozin, but not canagliflozin.

Table 2. Wound healing and amputation outcomes at 6 months according to SGLT-2 inhibitor use

	SGLT-2 inhibitors					Adjusted p
	No (n=102)	Yes (n=106)	Crude effect size (95% CI)	Crude p	Adjusted effect size (95% CI)	
Wound healing, n (%)						
A B C	37 (36.3%) 21 (20.6) 44 (43.1)	50 (47.2) 21 (19.8) 35 (33.0)	Common OR 0.64 (0.39-1.07)	0.090	aOR 0.77 (0.45-1.30)	0.324
Amputation, n (%)						
No Yes	72 (70.6) 30 (29.4)	83 (78.3) 23 (21.7)	OR 0.67 (0.35-1.25)	0.264	aOR 0.81 (0.42-1.55)	0.522

A: Improvement in Wagner grade, B: No change in Wagner grade, C: Progression in Wagner grade. Categorical variables were compared according to SGLT-2 inhibitor use using the chi-square test. For wound healing, the common odds ratio was estimated using ordinal logistic regression, with categories ordered as improvement, no change, and progression. Adjusted analyses included serum albumin, sodium, and spot urine albumin-to-creatinine ratio as covariates
OR: Odds ratio, aOR: Adjusted odds ratio, CI: Confidence interval, SGLT-2 inhibitors: Sodium-glucose cotransporter-2 inhibitors

Furthermore, the follow-up time for wound healing was similar. This may have contributed to the consistency of the findings between the two studies. In that study, wound care and antibiotic treatment were standardized for the patients (7). However, no such standardization was applied in our study.

As shown in the studies by Chen et al. (8) and Pavithra and Kaliyaperumal (9), the healing process of DFUs is influenced by numerous factors, including debridement, wound care, and antibiotic treatment. In the present study, the lack of standardization in antibiotic treatment and wound care may have had a significant impact on the results.

A study by Sung et al. (10) did not observe a clear association between the use of SGLT-2 inhibitors and an increased risk of amputation. This study, similar to our own, examined patients with active diabetic ulcers. The duration of DM and the HbA1c levels were comparable between the groups. Of the 27 patients using SGLT-2 inhibitors, only one used canagliflozin, while the others used empagliflozin or dapagliflozin. The similarity between the findings of the two studies may be explained by these factors. However, unlike our study, that study did not examine the wound healing process and the wound stages of the patients were not clearly determined.

Jiang et al. (11) associated peripheral artery disease and advanced Wagner's stage ulcers with increased amputation rates in patients with DFUs. The present study did not include patients with advanced peripheral artery disease or advanced Wagner's stage ulcers. Therefore, the findings of this study do not permit definitive conclusions regarding the overall safety of SGLT-2 inhibitor use in patients with DFUs.

Heyward et al. (12) reported no clear association between SGLT-2 inhibitor use and the overall risk of amputation. However, the amputation signal observed with canagliflozin requires further investigation. The CANVAS program (Canagliflozin Cardiovascular Assessment Study) revealed that patients treated with canagliflozin had approximately twice the risk of amputation (5). However, large randomized controlled trials, such as the EMPA-REG OUTCOME (Empagliflozin Cardiovascular Outcome Event Trial) and DECLARE-TIMI 58 (Dapagliflozin Effect on Cardiovascular Events) trials, did not find a similar increase in amputation risk for empagliflozin and dapagliflozin (13,14). Even in larger studies, the available evidence remains insufficient to draw definitive conclusions regarding the effect of SGLT-2 inhibitors on DFU healing or amputation risk.

Study Limitations

The interpretation of the findings of the present study should be considered in light of several limitations. First, the study was a single-center study with a relatively small sample size, which may limit the generalizability of our findings to other patient populations and clinical settings. Second, only patients with Wagner stage 1-2 ulcers were included; therefore, the findings cannot be generalized to patients with advanced ulcers or severe peripheral arterial disease. Third, wound care practices and antibiotic regimens were not completely standardized and may have affected wound healing outcomes, representing a potential source of confounding. Fourth, primary care doctors stopped SGLT-2 inhibitor treatment for some patients whose wound stage worsened during follow-up, but these patients were still observed and included in the analysis. In addition, no a priori sample size calculation was performed, as all eligible patients presenting during the predefined study period were consecutively included. Therefore, the study may have had limited statistical power to detect clinically meaningful between-group differences, and the possibility of a type II error cannot be excluded. Finally, insulin regimens and antidiabetic drug doses were not analyzed in detail. Despite these limitations, the prospective observational design, the inclusion of a well-defined population with Wagner stage 1-2 DFUs, and the evaluation of both wound-healing and amputation outcomes during a 6-month follow-up represent strengths

Table 3. Demographic, clinical, and laboratory characteristics of patients receiving SGLT-2 inhibitors according to treatment type			
	SGLT-2 inhibitor type		p
	Dapagliflozin (n=61)	Empagliflozin (n=45)	
Age, years	62.0 (36.0-76.0)	59.0 (38.0-71.0)	0.060
Sex, n (%)			
Male Female	43 (70.5) 18 (29.5)	32 (71.1) 13 (28.9)	0.945
Wagner grade, n (%)			
Grade 1 Grade 2	24 (39.3) 37 (60.7)	23 (51.1) 22 (48.9)	0.314
BMI (kg/m^2)	27.1 (22.7-37.8)	27.0 (23.6-35.2)	0.638
ALT (U/L)	14.0 (4.0-57.0)	16.0 (6.0-61.0)	0.067
AST (U/L)	16.0 (8.0-52.0)	17.0 (10.0-57.0)	0.107
Hemoglobin (g/dL)	11.8±1.8	12.4±1.9	0.118
WBC (×10^3/μL)	9.1 (3.9-20.9)	8.6 (4.8-33.0)	0.964
Neutrophil count (×10^3/μL)	6.2 (2.2-19.9)	5.7 (2.3-30.0)	0.906
Lymphocyte count (×10^3/μL)	1.8±0.8	2.1±0.7	0.057
Monocyte count (×10^3/μL)	0.6 (0.3-1.3)	0.7 (0.1-1.8)	0.524
Platelet count (×10^3/μL)	279.0 (144.0-690.0)	298.0 (140.0-525.0)	0.742
Serum albumin (g/dL)	4.0 (1.6-4.7)	3.9 (2.5-4.4)	0.529
eGFR (mL/min/1.73 m^2)	91.0 (38.0-132.0)	88.0 (41.0-120.0)	0.992
eGFR			
≥90 mL/min/1.73 m^2 <90 mL/min/1.73 m^2	31 (50.8) 30 (49.2)	22 (48.9) 23 (51.1)	0.844
Urea (mg/dL)	41.9±18.4	42.6±18.1	0.853
Creatinine (mg/dL)	0.9 (0.4-2.2)	0.9 (0.4-1.6)	0.974
C-reactive protein (mg/L)	40.0 (4.0-270.0)	23.0 (1.0-234.0)	0.130
Sodium (mmol/L)	138.0 (127.0-145.0)	139.0 (130.0-141.0)	0.775
Potassium (mmol/L)	4.5±0.5	4.5±0.5	0.741
HbA1c-baseline	9.4 (6.3-14.7)	9.6 (6.8-18.2)	0.383
HbA1c-6th month	9.0 (5.8-14.2)	8.9 (6.0-17.0)	0.913
Duration of diabetes, years	8.0 (2.0-20.0)	8.0 (1.0-16.0)	0.430
Spot urine albumin-to-creatinine ratio (mg/g)	84.0 (5.0-4239.0)	36.0 (2.0-3510.0)	0.106
Hyperbaric oxygen therapy, n (%)			
No Yes	55 (90.2) 6 (9.8)	37 (82.2) 8 (17.8)	0.366

Unless otherwise indicated, variables are expressed as mean ± standard deviation or median (minimum-maximum). Categorical variables were compared according to SGLT-2 inhibitor use using the chi-square test. Continuous variables are presented as mean ± standard deviation and were compared using the independent-samples t-test, or as median (min-max) and were compared using the Mann-Whitney U test, as appropriate

BMI: Body mass index, ALT: Alanine aminotransferase, AST: Aspartate aminotransferase, WBC: White blood cell count, eGFR: Estimated glomerular filtration rate, HbA1c: Glycated hemoglobin, SGLT-2 inhibitors: Sodium-glucose cotransporter-2 inhibitors

of the present study. In addition, multivariable-adjusted analyses accounted for potential baseline confounders.

Conclusion

This study did not find a clear association between use of dapagliflozin or empagliflozin and impaired wound healing or increased amputation rates in patients with Wagner stage 1-2 DFUs. However, the findings should be interpreted with caution, given the observational design and the limited statistical power of the study. Larger-scale prospective studies with sufficient statistical power are required to establish a clearer association in this regard.

Ethics

Ethics Committee Approval: The study was approved by the Clinical Research Ethics Committee of University of

Table 4. Wound healing and amputation outcomes at 6 months according to SGLT-2 inhibitor type			
	SGLT-2 inhibitor type		p
	Dapagliflozin (n=61)	Empagliflozin (n=45)	
Wound healing, n (%)			
A B C	27 (44.2) 10 (16.4) 24 (39.3)	23 (51.1) 11 (24.4) 11 (24.4)	0.241
Amputation, n (%)			
No Yes	44 (72.1) 17 (27.9)	39 (86.7) 6 (13.3)	0.120

A: Improvement in Wagner grade, B: No change in Wagner grade, C: Progression in Wagner grade. Categorical variables were compared according to SGLT-2 inhibitor type using the chi-square test
SGLT-2 inhibitors: Sodium-glucose cotransporter-2 inhibitors

Health Sciences Türkiye, Adana City Training and Research Hospital (decision no: 2372, date: 12.01.2023).

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

Footnotes

Authorship Contributions

Surgical and Medical Practices: Z.G.T., T.S., Concept: Z.G.T., T.S., Design: Z.G.T., S.P.Y.K., T.S., Data Collection or Processing: Z.G.T., S.P.Y.K., T.S., Analysis or Interpretation: Z.G.T., S.P.Y.K., T.S., Literature Search: Z.G.T., T.S., Writing: Z.G.T., S.P.Y.K., T.S.

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

References

1. Harreiter J, Roden M. Diabetes mellitus—Definition, klassifikation, diagnose, screening und prävention (update 2026) [Diabetes mellitus-definition, classification, diagnosis, screening and prevention (update 2026)]. Wien Klin Wochenschr. 2026;138(Suppl 4):76-86. German.
2. Kim JM, Kang SM, Jung JH, et al. Comprehensive management of diabetic foot ulcers: current evidence and practice. J Diabetes. 2026;18:e70221.
3. American Diabetes Association Professional Practice Committee for Diabetes*. 12. Retinopathy, neuropathy, and foot care: standards of care in diabetes-2026. Diabetes Care. 2026;49:261-76.
4. American Diabetes Association Professional Practice Committee for Diabetes*. 9. Pharmacologic approaches to glycemic treatment: standards of care in diabetes-2026. Diabetes Care. 2026;49:183-215.
5. Perkovic V, de Zeeuw D, Mahaffey KW, et al. Canagliflozin and renal outcomes in type 2 diabetes: results from the CANVAS Program randomised clinical trials. Lancet Diabetes Endocrinol. 2018;6:691-704.
6. Lin C, Zhu X, Cai X, et al. SGLT2 inhibitors and lower limb complications: an updated meta-analysis. Cardiovasc Diabetol. 2021;20:91.
7. Dumortier C, Aho Glele S, Rouland A, et al. Effects of sodium-glucose cotransporter 2 inhibitor use on mortality, amputation, and healing in patients with diabetic foot ulcer. Diabetes Care. 2026;49:674-81.
8. Chen P, Ahluwalia R, Chuo CB, Johnson-Lynn SE. Management of the acute diabetic foot: an update. Bone Joint J. 2026;108-B:322-30.
9. Pavithra TR, Kaliyaperumal R. Epidemiology, pathophysiology, and management of diabetic foot: a comprehensive review. Foot (Edinb). 2026;66:102225.
10. Sung J, Padmanabhan S, Gurung S, et al. SGLT2 inhibitors and amputation risk: real-world data from a diabetes foot wound clinic. J Clin Transl Endocrinol. 2018;13:46-7.
11. Jiang Y, Ran X, Jia L, et al. Epidemiology of type 2 diabetic foot problems and predictive factors for amputation in China. Int J Low Extrem Wounds. 2015;14:19-27.
12. Heyward J, Mansour O, Olson L, Singh S, Alexander GC. Association between sodium-glucose cotransporter 2 (SGLT2) inhibitors and lower extremity amputation: a systematic review and meta-analysis. PLoS One. 2020;15:e0234065.
13. Zinman B, Wanner C, Lachin JM, et al. Empagliflozin, cardiovascular outcomes, and mortality in type 2 diabetes. N Engl J Med. 2015;373:2117-28.
14. Mosenzon O, Wiviott SD, Cahn A, et al. Effects of dapagliflozin on development and progression of kidney disease in patients with type 2 diabetes: an analysis from the DECLARE-TIMI 58 randomised trial. Lancet Diabetes Endocrinol. 2019;7:606-17.

Original Article

DOI: 10.4274/haseki.galenos.2026.46036
Med Bull Haseki 2026;64(5):391-398

Anxiety, Depression, and Sleep Disturbances in Patients with Liver Cirrhosis: Associations with Disease Severity

Zeynep Busra Akin[1], Mustafa Can Senoymak[1], Alpaslan Tanoglu[2], Suleyman Bas[1], Murat Yeniceri[3]

[1]*University of Health Sciences Türkiye, Sancaktepe Sehit Prof. Dr. Ilhan Varank Training and Research Hospital, Clinic of Internal Medicine, Istanbul, Türkiye*

[2]*Bahcesehir University Faculty of Medicine, Department of Internal Medicine, Gastroenterology, Istanbul, Türkiye*

[3]*University of Health Sciences Türkiye, Kartal Dr. Lutfi Kirdar City Hospital, Clinic of Rheumatology, Istanbul, Türkiye*

Abstract

Aim: Depression and sleep disturbances are common but often overlooked complications in patients with liver cirrhosis. We aimed to determine the prevalence of anxiety, depression, and sleep disturbances in patients with liver cirrhosis and to investigate their association with disease severity.

Methods: This cross-sectional study enrolled 115 patients with cirrhosis and 207 healthy volunteers, all recruited between October 2022 and June 2023. Sociodemographic characteristics (age, gender, marital status, smoking and alcohol use, income, and education level), medical history, and laboratory findings were recorded. Disease severity was evaluated using the Child-Pugh and MELD-Na scoring systems. The Hospital Anxiety and Depression Scale (HADS) and the Pittsburgh Sleep Quality Index (PSQI) were used to assess anxiety, depression, and sleep disturbances. Statistical comparisons were performed using the Mann-Whitney U test, the chi-square test, and the Spearman correlation analysis.

Results: Patients with liver cirrhosis had significantly higher median HADS-Anxiety scores [13 (8-17) vs. 5 (4-7), p<0.001], HADS-Depression scores [12 (9-16) vs. 4 (3-6), p<0.001], and PSQI scores [11 (6-14) vs. 5 (4-7), p<0.001] compared with controls. Anxiety was present in 64.3% of patients versus 3.4% of controls (p<0.001); depression in 66% of patients versus 1.9% of controls (p<0.001); and poor sleep quality in 85.2% of patients versus 47.8% of controls (p<0.001). Significant positive correlations were observed between HADS-Anxiety and MELD-Na scores and between HADS-Depression and PSQI scores and both MELD-Na and Child-Pugh scores. No statistically significant correlation was found between HADS-Anxiety and Child-Pugh scores (p=0.056).

Conclusion: Anxiety, depression, and sleep disturbances are highly prevalent among patients with cirrhosis and are associated with disease severity. Routine psychological assessment and timely intervention may improve patients' quality of life and clinical outcomes.

Keywords: Cirrhosis, depression, sleep-wake disorders, anxiety

Introduction

Liver cirrhosis is a chronic and progressive condition that represents the final stage of various liver diseases, including chronic alcohol use, viral hepatitis, autoimmune liver disorders, and other etiologies. It is characterized by persistent liver injury and fibrotic tissue formation, which disrupt normal liver function. Cirrhosis is associated with high morbidity and mortality worldwide, and patients experience frequent hospitalizations, polypharmacy, invasive procedures, and a substantial reduction in quality of life (1). Psychological comorbidities, particularly anxiety

Corresponding Author: Murat Yeniceri, MD, University of Health Sciences Türkiye, Kartal Dr. Lutfi Kirdar City Hospital, Clinic of Rheumatology, Istanbul, Türkiye
E-mail: 4256.gatf@gmail.com **ORCID:** orcid.org/0000-0003-3046-5784
Received: 17.03.2026 **Accepted:** 23.08.2026 **Publication Date:** 17.11.2026

Cite this article as: Akin ZB, Senoymak MC, Tanoglu A, Bas S, Yeniceri M. Anxiety, depression, and sleep disturbances in patients with liver cirrhosis: associations with disease severity. Med Bull Haseki. 2026;64(5):391-398

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

and depression, are common in patients with cirrhosis and may contribute to disease progression and adverse clinical outcomes (2,3). Recent evidence highlights that these psychiatric comorbidities significantly influence disease burden and patient-reported outcomes in cirrhosis (4). Sleep disturbances are also prevalent among patients with cirrhosis, further contributing to impaired daily functioning and reduced overall well-being (5,6). Recent studies indicate that cirrhosis may directly impair sleep quality through complex pathophysiological mechanisms and is associated with worsened prognosis and quality of life (7).

Although the relationship between liver cirrhosis and depression is well recognized, few studies have simultaneously assessed anxiety, depression, and sleep disturbances in this population. Previous studies have primarily focused on individual symptoms or specific etiologies such as hepatitis B-related cirrhosis (8-10). Comprehensive investigations examining the interaction between these psychological factors and disease severity, including MELD-Na and Child-Pugh scores, are limited. Recent studies suggest that disease severity indices such as MELD-Na are associated with both physical and psychological outcomes in cirrhosis (11). We hypothesized that patients with liver cirrhosis would have higher rates of depression, anxiety, and sleep disturbances than healthy individuals, and that these conditions would correlate with liver disease severity.

While several studies have addressed psychological symptoms in chronic liver disease, most have examined either depression or sleep disturbances. Studies that evaluate anxiety, depression, and sleep quality together and their association with disease severity in patients with cirrhosis remain scarce. Therefore, further studies are warranted to clarify the combined impact of these psychological and behavioral factors on the clinical course of cirrhosis. Emerging evidence also emphasizes the role of psychological distress and resilience in shaping health-related quality of life in patients with liver cirrhosis (12). A recent systematic review and meta-analysis also confirmed that anxiety and depression are highly prevalent among patients with liver cirrhosis, emphasizing the importance of routine psychological assessment in this population (13). Accordingly, this study was designed to assess the prevalence of anxiety, depression, and sleep disturbances in patients with liver cirrhosis and to investigate their association with disease severity. Understanding these associations may facilitate the early identification of patients at increased psychological risk and support more comprehensive clinical assessment.

Materials and Methods

Compliance with Ethical Standards

Ethical approval for this study was obtained from the Clinical Research Ethics Committee of the University of Health Sciences Türkiye, Sancaktepe Sehit Prof. Dr. Ilhan Varank Training and Research Hospital (approval date: September 14, 2022; approval no: 114).

Study Design and Participants

This study was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to inclusion in the study.

A total of 115 adult patients with liver cirrhosis and 207 healthy volunteers were enrolled in the study. Cirrhosis was confirmed based on clinical, laboratory, histological, and radiological findings. All participants received detailed information about the study and provided written informed consent. Sociodemographic data, including age, gender, marital status, smoking and alcohol habits, income level, and educational level, were recorded.

Inclusion and Exclusion Criteria

Adult patients with confirmed liver cirrhosis and healthy volunteers without known chronic diseases or substance use were included in the study. Participants with cerebrovascular disease or cerebral trauma, with recent hepatic encephalopathy or receiving treatment for hepatic encephalopathy, with psychiatric comorbidities, with known sleep disorders, who used central nervous system-acting medications, who were unable to complete the questionnaires, or who refused to provide informed consent were excluded.

Assessments

Clinical and Laboratory Evaluation: Routine clinical examinations, medical history, medication use, and laboratory parameters were recorded. Disease severity was assessed using the Child-Pugh classification and the MELD-Na score.

Psychological Assessment: Anxiety and depression were assessed using the Hospital Anxiety and Depression Scale (HADS). HADS scores were categorized as ≤7 (normal), 8-10 (borderline), and ≥11 (abnormal). Sleep quality was evaluated using the Pittsburgh Sleep Quality Index (PSQI). A global PSQI score ≥5 indicated poor sleep quality.

The overall study design, participant selection process, clinical and psychological assessments, and statistical analyses are summarized in Figure 1.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics for Windows, version 25.0 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean ±

standard deviation (SD) or median [interquartile range (IQR)], as indicated in the tables. Categorical variables are presented as counts and percentages. Independent-samples t-tests were used to compare age between patients with cirrhosis and healthy controls and to compare HADS-Anxiety, HADS-Depression, PSQI, MELD-Na, and Child-Pugh scores between patients with decompensated and compensated cirrhosis. The Mann-Whitney U test was used to compare HADS-Anxiety, HADS-Depression, and PSQI scores between patients with cirrhosis and healthy controls. Categorical variables were compared using the chi-square test. Correlations were assessed using Spearman's rank correlation. A two-sided p-value <0.05 was considered statistically significant.

Results

A total of 322 participants were included in the study, comprising 115 patients with liver cirrhosis and 207 healthy controls. No significant differences between the groups were observed with respect to age, sex, marital status, income level, or educational level (all p>0.05, Table 1).

Patients with liver cirrhosis had significantly higher HADS-Anxiety [median 13 (8-17) vs. 5 (4-7), p<0.001], HADS-Depression [median 12 (9-16) vs. 4 (3-6), p<0.001] and PSQI scores [median 11 (6-14) vs. 5 (3-7), p<0.001] than healthy controls (Table 2, Figures 1 and 2). The prevalences of anxiety, depression, and poor sleep quality were significantly higher in the cirrhotic group (64.3%, 66%, and 85.2%, respectively) than in the control group (3.4%, 1.9%, and 47.8%; p<0.001 for all, Table 3 and Figure 3).

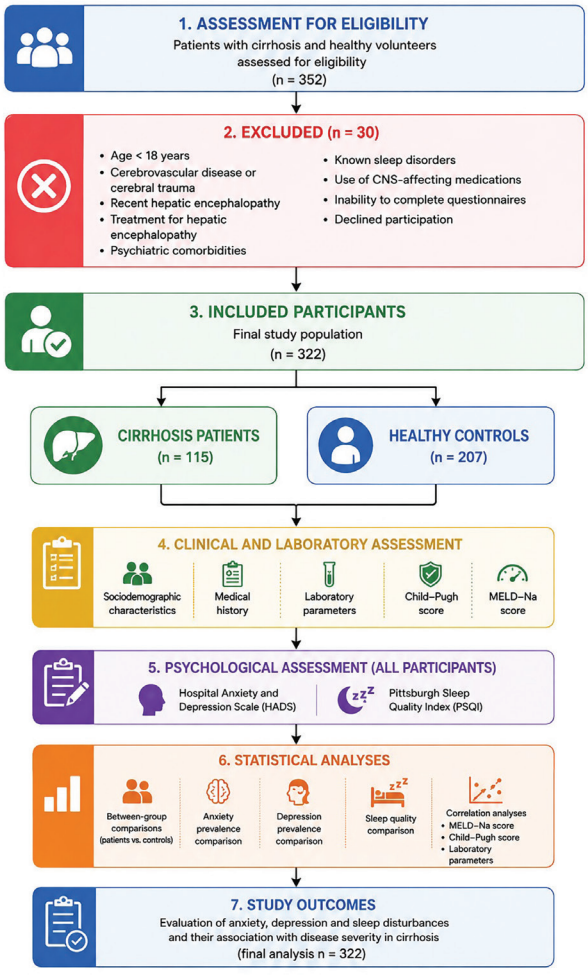


Figure 1. Flow diagram of the study design, participant selection, assessments and statistical analyses
HADS: Hospital Anxiety and Depression Scale, PSQI: Pittsburgh Sleep Quality Index, MELD-Na: Model for End-Stage Liver Disease-Sodium

Table 1. Comparison of sociodemographic data of patients with liver cirrhosis and control groups

Variable	Cirrhosis group (n=115)	Healthy controls (n=207)	p
Age (year)	62.42±12.00[a]	60.02±13.70[a]	0.051[b]
Sex			
Male (n=168)	62 (36.90%)	106 (63.10%)	0.641[c]
Female (n=154)	53 (34.40%)	101 (65.60%)	
Income level			
Below minimum wage (n=91)	33 (35.60%)	58 (64.40%)	0.622[c]
Minimum wage (n=98)	38 (38.80%)	60 (61.20%)	
Above minimum wage (n=133)	44 (32.60%)	89 (67.40%)	
Marital status			
Single (n=129)	51 (44.30%)	78 (37.70%)	0.242[c]
Married (n=193)	64 (55.70%)	129 (62.60%)	
Education level			
Illiterate (n=52)	19 (16.50%)	33 (15.90%)	0.131[c]
Primary school (n=87)	39 (34.80%)	48 (23.20%)	
High school (n=90)	29 (24.30%)	61 (29.50%)	
University (n=93)	28 (24.30%)	65 (31.40%)	

[a]: Mean ± standard deviation, [b]: Independent-samples t-test, [c]: Chi-square test

Table 2. Comparison of HADS-Anxiety, HADS-Depression and Pittsburgh Sleep Quality Index scores between the cirrhotic patient group and the control group

Variable	Cirrhosis group (n=115)	Healthy controls (n=207)	p
HADS-Anxiety score	13 (8-17)[a]	5 (4-7)[a]	<0.001[b]
HADS-Depression score	12 (9-16)[a]	4 (3-6)[a]	<0.001[b]
Pittsburgh Sleep Quality Index score	11 (6-14)[a]	4 (3-6)[a]	<0.001[b]

[a]: Mean ± standard deviation, [b]: Independent-samples t-test, HADS: Hospital Anxiety and Depression Scale

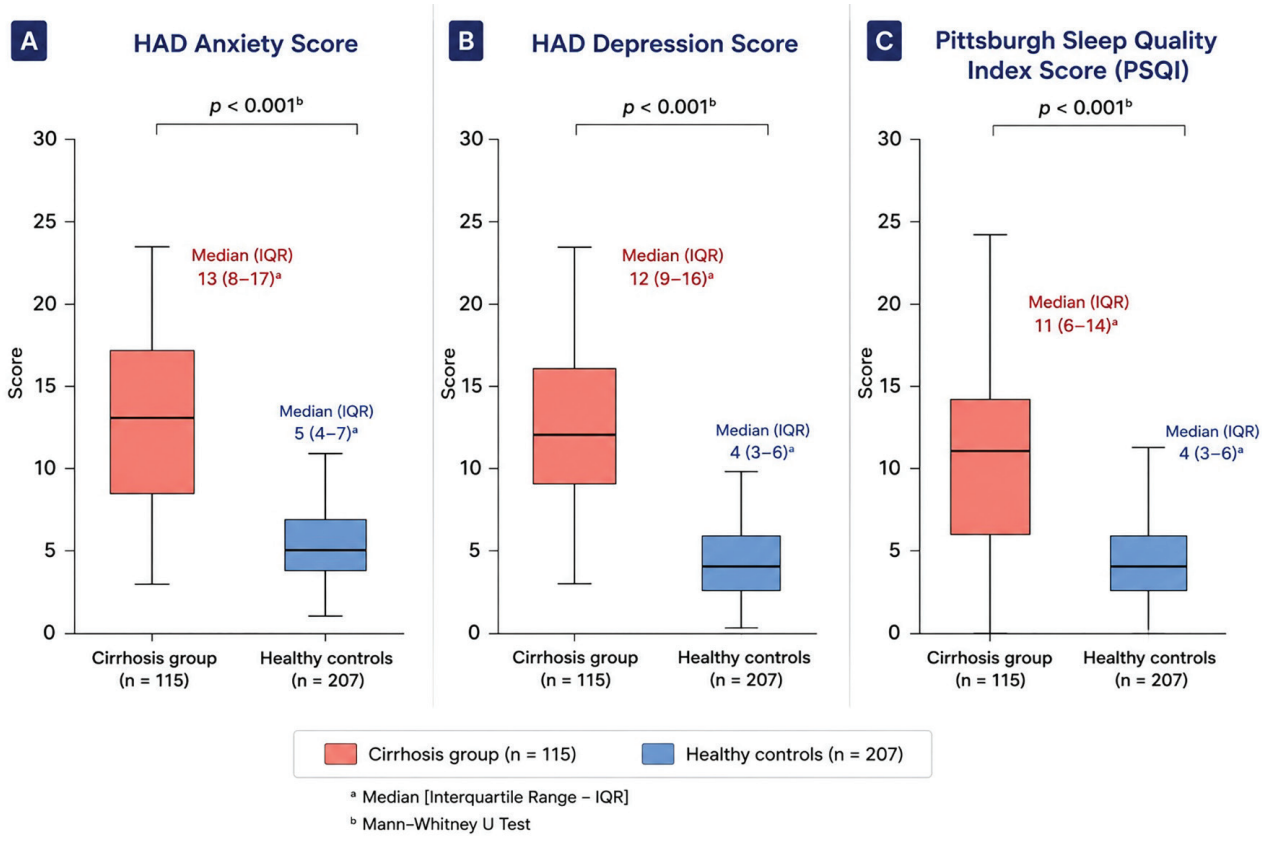


Figure 2. Comparison of HADS-Anxiety, HADS-Depression and Pittsburgh Sleep Quality Index (PSQI) scores between patients with cirrhosis and healthy controls. (A) HADS-Anxiety scores; (B) HADS-Depression scores; (C) PSQI scores. Data are presented as median IQR

HADS: Hospital Anxiety and Depression Scale; PSQI: Pittsburgh Sleep Quality Index; IQR: Interquartile range.

Table 3. Comparison of the cirrhotic patient group and the control group according to HADS-Anxiety, HADS-Depression and Pittsburgh Sleep Quality Index scores

Variable	Cirrhosis group (n=115)	Healthy controls (n=207)	p[a]
HADS-Anxiety score			
No anxiety (n=187, 58.1%)	19 (16.5%)	168 (81.2%)	**<0.001**
Borderline anxiety (n=54, 16.8%)	22 (19.1%)	32 (15.4%)	
Anxiety (n=81, 25.2%)	74 (64.3%)	7 (3.4%)	
HADS-Depression score			
No depression (n=205, 63.7%)	21 (18.3%)	184 (88.9%)	**<0.001**
Borderline depression (n=37, 11.5%)	18 (15.7%)	19 (9.2%)	
Depression (n=80, 24.8%)	76 (66.0%)	4 (1.9%)	
Pittsburgh Sleep Quality Index			
Good sleep quality (n=125, 38.8%)	17 (14.8%)	108 (52.2%)	**<0.001**
Poor sleep quality (n=197, 61.2%)	98 (85.2%)	99 (47.8%)	

p[a]: Chi-square test, HADS: Hospital Anxiety and Depression Scale

Table 4. Relationship between HADS-Anxiety score, HADS-Depression score, Pittsburgh Sleep Quality Index score and MELD-Na and Child-Pugh score in cirrhosis patient group

	Cirrhosis patients	
	Rho p[a]	
HADS-Anxiety, HADS-Depression	0.754	**0.001**
HADS-anxiety, Pittsburgh Sleep Quality Index	0.405	**0.001**
HADS-anxiety, MELD-Na score	0.275	**0.004**
HADS-anxiety, Child-Pugh score	0.183	0.056
HADS-Depression, Pittsburgh Sleep Quality Index	0.553	**0.001**
HADS-Depression, MELD-Na score	0.359	**0.001**
HADS-Depression, Child-Pugh score	0.289	**0.002**
Pittsburgh Sleep Quality Index-MELD-Na score	0.285	**0.002**
Pittsburgh Sleep Quality Index-Child-Pugh score	0.262	**0.005**

rho: Correlation coefficient, [a]: Spearman Correlation test, HADS: Hospital Anxiety and Depression Scale, MELD-Na: Model for End-Stage Liver Disease-Sodium

Table 5. Comparison of HADS-Anxiety, HADS-Depression, Pittsburgh Sleep Quality Index, MELD-Na and Child scores between decompensated and compensated cirrhotic patient groups

	Patients with decompensated cirrhosis (n=88)	Patients with compensated cirrhosis (n=27)	p
HADS-Anxiety	12.90±5.72[a]	12.70±5.30[a]	0.953[b]
HADS-Depression	12.29±5.21[a]	12.09±5.16[a]	0.968[b]
Pittsburgh Sleep Quality Index	10.36±4.70[a]	9.17±4.11[a]	0.335[b]
MELD-Na score	12.49±6.23[a]	10.08±6.20[a]	**0.021[b]**
Child score	7.21±1.74[a]	5.71±1.40[a]	**0.001[b]**

[a]: Mean ± standard deviation, [b]: Mann-Whitney U test, HADS: Hospital Anxiety and Depression Scale, MELD-Na: Model for End-Stage Liver Disease-Sodium

Significant positive correlations were observed between HADS-Anxiety scores and MELD-Na scores, as well as between HADS-Depression and PSQI scores and both disease severity indices (MELD-Na and Child-Pugh scores). However, no statistically significant correlation was detected between HADS-Anxiety and Child-Pugh scores (r=0.183, p=0.056; Table 4). No significant differences were found in HADS or PSQI scores between compensated and decompensated cirrhosis; however, MELD-Na and Child-Pugh scores were higher in decompensated patients (p=0.021 and p=0.001, Table 5). Correlation analyses further demonstrated significant associations between psychological assessment scores and selected laboratory parameters, including INR, bilirubin, albumin, sodium, and CRP levels (Table 6).

Discussion

This study demonstrated that patients with liver cirrhosis had significantly higher levels of anxiety and depression and poorer sleep quality than healthy controls. Our findings further showed that the associations between psychological symptoms and disease severity varied according to the assessment tool and severity index. While HADS-Anxiety scores were significantly correlated with MELD-Na scores, no statistically significant correlation was found with Child-Pugh scores. In contrast, HADS-Depression and PSQI scores were significantly correlated with both MELD-Na and Child-Pugh scores, highlighting the potential clinical importance of psychological assessment in the management of liver cirrhosis.

The high prevalence of anxiety and depression observed in our cohort suggests that psychological distress is a major yet frequently underrecognized component of liver cirrhosis. Our findings are consistent with the recent systematic review and meta-analysis by Abureesh et al. (13), which confirmed that anxiety and depression are highly prevalent among patients with liver cirrhosis. Unlike earlier studies that primarily evaluated depression or sleep disturbances separately, our study simultaneously assessed anxiety, depression, and sleep quality, providing a more comprehensive evaluation of the psychological burden associated with liver cirrhosis.

The high prevalence of poor sleep quality in our cohort may be explained by several pathophysiological mechanisms associated with liver cirrhosis, including circadian rhythm disruption, altered melatonin metabolism, hepatic encephalopathy, and systemic inflammation (14-17). These mechanisms may also contribute to the close relationship between sleep disturbances and mood disorders observed in our study (17).

Table 6. Relationship between HADS-Anxiety, HADS-Depression and PSQI scores and biochemical and hematological parameters in patients with cirrhosis

	HADS-Anxiety score		Pittsburgh Sleep Quality Index		HADS-Depression score	
	r	p[a]	r	p[a]	r	p[a]
Urea	0.054	0.576	0.093	0.334	0.061	0.525
Creatinine	-0.057	0.554	0.090	0.343	0.054	0.567
AST	0.113	0.234	0.043	0.655	0.064	0.499
ALT	-0.044	0.645	-0.113	0.234	-0.087	0.359
Direct bilirubin	0.136	0.158	0.141	0.138	**0.204**	**0.030**
Total bilirubin	0.179	0.059	0.056	0.562	**0.192**	**0.043**
Albumin	-0.122	0.205	**-0.210**	**0.028**	**-0.286**	**0.002**
Sodium	**-0.309**	**0.001**	**-0.249**	**0.008**	**-0.310**	**0.001**
Potassium	-0.014	0.884	-0.058	0.544	-0.104	0.274
LDL-cholesterol	-0.104	0.372	**-0.244**	**0.033**	-0.222	0.052
HDL-cholesterol	0.058	0.620	**-0.233**	**0.041**	-0.138	0.231
Total cholesterol	-0.072	0.530	**-0.326**	**0.003**	**-0.251**	**0.025**
Triglyceride	-0.007	0.950	0.015	0.898	-0.053	0.639
Leukocyte	**0.222**	**0.019**	0.064	0.500	0.155	0.101
Hemoglobin	-0.048	0.612	**-0.291**	**0.002**	-0.093	0.328
Platelet	0.007	0.946	-0.108	0.254	-0.048	0.616
Vitamin B12	0.097	0.393	0.115	0.308	0.002	0.985
INR	**0.203**	**0.036**	**0.286**	**0.003**	**0.253**	**0.008**
CRP	0.123	0.195	**0.222**	**0.018**	**0.198**	**0.035**
CRP/ALB ratio	0.139	0.293	**0.323**	**0.012**	**0.358**	**0.005**

r: Correlation Coefficient, [a]: Spearman Correlation test, HADS: Hospital Anxiety and Depression Scale, PSQI: Pittsburgh Sleep Quality Index, AST: Aspartate aminotransferase, ALT: Alanine aminotransferase, LDL: Low-density lipoprotein, HDL: High-density lipoprotein, INR: International normalized ratio, CRP: C-reactive protein, ALB: Albumin

Our findings further demonstrated that the associations between psychological symptoms and disease severity varied according to the assessment tool and severity index. Higher HADS-Anxiety scores were significantly correlated with higher MELD-Na scores, whereas no statistically significant correlation was observed with Child-Pugh scores. In contrast, higher HADS-Depression and PSQI scores were significantly correlated with both MELD-Na and Child-Pugh scores. These findings highlight the importance of routine psychological assessment and intervention in patients with liver cirrhosis. Similar associations have also been reported in previous studies (18-20). These findings are also consistent with recent evidence emphasizing the importance of recognizing and addressing psychological symptoms in patients with advanced liver disease (21).

The main strength of our study is its comprehensive evaluation of anxiety, depression, and sleep quality across all etiologies of cirrhosis, along with their associations with MELD-Na and Child-Pugh classifications. To the best of our knowledge, this is the first study from Turkey to simultaneously evaluate anxiety, depression, sleep quality, and their associations with both MELD-Na and Child-Pugh scores in a single cohort of patients with liver cirrhosis.

Study Limitations

This study has several limitations, including its relatively small sample size, single-center design, and the lack of a detailed assessment of medication effects and psychosocial factors that may influence psychological symptoms and sleep quality. In addition, because the study was conducted at a tertiary referral center, the population may have included patients with more advanced or complex liver disease, potentially introducing referral bias and limiting the generalizability of our findings to the broader population with liver cirrhosis. Despite these limitations, our findings provide clinically relevant evidence regarding the psychological burden associated with liver cirrhosis and underscore the importance of routine psychological assessment in clinical practice.

Conclusion

Patients with liver cirrhosis exhibited significantly higher levels of anxiety and depression, as well as poorer sleep quality than healthy controls. These psychological symptoms were significantly associated with disease severity, as reflected by MELD-Na and Child-Pugh scores. Our findings support incorporating routine psychological assessment

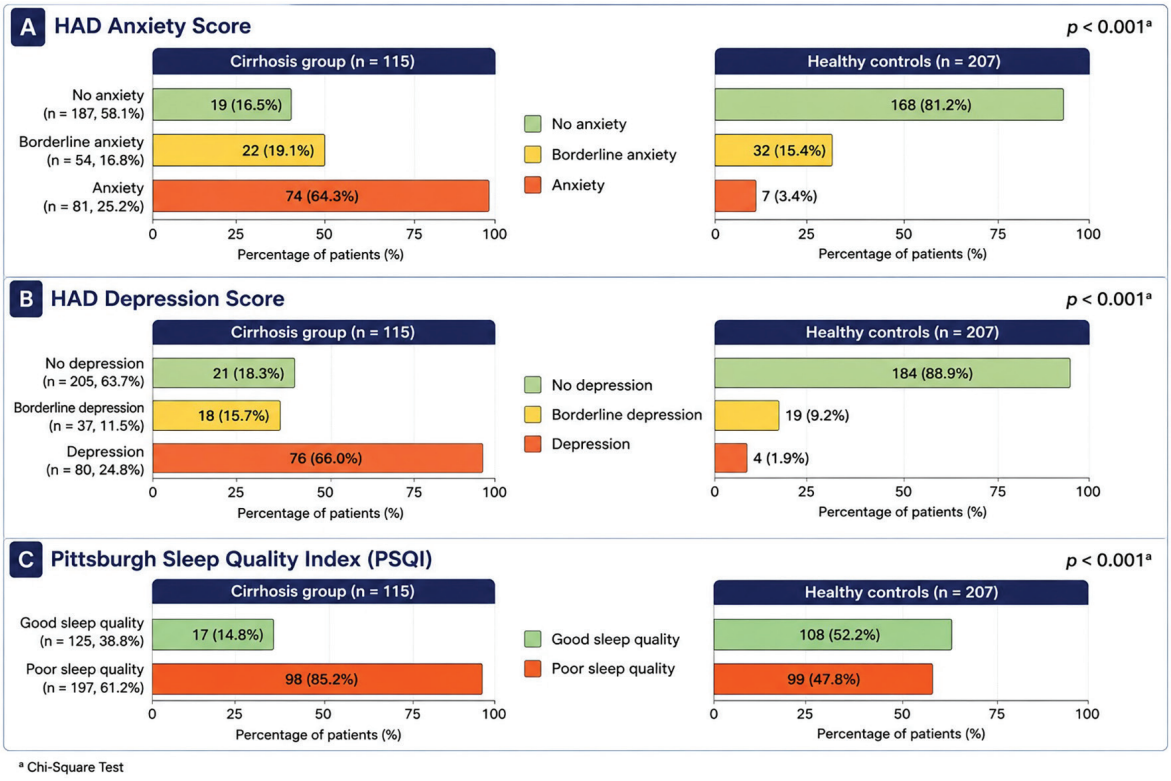


Figure 3. Distribution of anxiety, depression and sleep quality categories in patients with cirrhosis and healthy controls. (A) Anxiety categories according to HADS-Anxiety scores (no anxiety, borderline anxiety, and anxiety); (B) depression categories according to HADS-Depression scores (no depression, borderline depression, and depression); (C) sleep quality categories according to Pittsburgh Sleep Quality Index (PSQI) scores (good and poor sleep quality). Bars represent the number and percentage of participants in each category. Between-group comparisons were performed using the chi-square test (p<0.001 for all comparisons)

into the comprehensive evaluation of patients with liver cirrhosis, particularly those with more advanced disease. Early recognition of psychological symptoms may facilitate timely referral for appropriate mental health assessment and multidisciplinary care. Further multicenter prospective studies are warranted to confirm these findings.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the Clinical Research Ethics Committee of the University of Health Sciences Türkiye, Sancaktepe Sehit Prof. Dr. Ilhan Varank Training and Research Hospital (approval date: September 14, 2022; approval no: 114).

Informed Consent: This study was conducted in accordance with the principles of the Declaration of Helsinki. Written informed consent was obtained from all participants prior to inclusion in the study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Z.B.A., Concept: Z.B.A., Design: Z.B.A., A.T., M.C.S., Data Collection or Processing: M.Y., Analysis or Interpretation: A.T., S.B., Literature Search: Z.B.A., M.C.S., Writing: A.T., S.B.

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

Declaration Regarding the Use of AI and AI-Assisted Technologies: Artificial intelligence-assisted tools were used for language editing, manuscript formatting, table preparation, and literature checking. All outputs were reviewed and verified by the authors. The authors take full responsibility for the content of the manuscript, the accuracy of the data, and the integrity of the work.

References

1. Peng JK, Hepgul N, Higginson IJ, Gao W. Symptom prevalence and quality of life of patients with end-stage liver disease: a systematic review and meta-analysis. Palliat Med. 2019;33:24-36.
2. Qin S, Cheng X, Zhang S, et al. Sleep patterns, genetic susceptibility, and risk of cirrhosis among individuals with non-alcoholic fatty liver disease. Hepatol Int. 2024;18:1158-67.

3. Dong S, Zhou S, Liu J, et al. Associations between sleep disorders and clinical outcomes of patients with primary biliary cholangitis. Adv Med Sci. 2024;69:385-90.
4. Zimbrean PC, Lok AS, Weinrieb RM, et al. Depression and anxiety management in cirrhosis. Hepatol Commun. 2025;9:e1234.
5. Zhao M, Yan Y, Wang X, Liu B, Luo X. Sleep disturbance in patients with cirrhosis and transjugular intrahepatic portosystemic shunt. BMC Gastroenterol. 2024;24:381.
6. Cooper KM, Colletta A, Osorio B, Herringshaw E, Talat A, Devuni D. History of depression is associated with higher prevalence of hepatic encephalopathy in patients with advanced liver disease. Am J Med. 2024;137:872-9.
7. Feng J, Wang Y, Li X, et al. Sleep disorders in patients with liver cirrhosis: mechanisms and clinical implications. J Clin Transl Hepatol. 2025;13:345-54.
8. Abureesh M, Alkhayyat M, Abualnadi I, et al. Epidemiology of depressive disorders in patients with liver cirrhosis: a population-based study in the United States. Prim Care Companion CNS Disord. 2022;24:20m02889.
9. Hernaez R, Kramer JR, Khan A, et al. Depression and anxiety are common among patients with cirrhosis. Clin Gastroenterol Hepatol. 2022;20:194-203.
10. Xiao G, Ye Q, Han T, Yan J, Sun L, Wang F. Study of the sleep quality and psychological state of patients with hepatitis B liver cirrhosis. Hepatol Res. 2018;48:275-82.
11. Hansen L, Patel K, Nguyen M, et al. Longitudinal changes in physical and psychological outcomes in cirrhosis. Clin Gastroenterol Hepatol. 2025;23:789-98.
12. Qiu S, Zhang H, Li Y, et al. Resilience and health-related quality of life in patients with cirrhosis. Sci Rep. 2026;16:12345.
13. Abureesh O, Al-Shraideh A, Sleiman J, et al. Global prevalence and predictors of depression and anxiety in patients with liver cirrhosis: a systematic review and meta-analysis. Front Epidemiol. 2026;6:1740227.
14. Bruyneel M, Sersté T. Sleep disturbances in patients with liver cirrhosis: prevalence, impact, and management challenges. Nat Sci Sleep. 2018;10:369-75.
15. Nardelli S, Pentassuglio I, Pasquale C, et al. Depression, anxiety and alexithymia symptoms are major determinants of health related quality of life (HRQoL) in cirrhotic patients. Metab Brain Dis. 2013;28:239-43.
16. Montagnese S, De Pittà C, De Rui M, et al. Sleep-wake abnormalities in patients with cirrhosis. Hepatology. 2014;59:705-12.
17. Zhao M, Yan Y, Wang X, et al. Sleep disturbance in patients with cirrhosis and transjugular intrahepatic portosystemic shunt. BMC Gastroenterol. 2024;24:38
18. Yao S, Peng B, Liu W, Yang J, Hui Y, Sun C. Skeletal muscle abnormalities negatively linked to sleep disturbances in patients with liver cirrhosis: a cross-sectional study. Eur J Med Res. 2025;30:917.
19. Zhang T, Liu H, Li L, Wang Y, Chen Y. Anxiety, depression, and sleep disorders in patients with hepatitis B-related cirrhosis: a cross-sectional study. Medicine(Baltimore). 2017;96:e7967.
20. Jeong SW, Jang JY, Kim YJ, et al. Clinical significance of depression and anxiety in patients with liver cirrhosis. Clinical and Molecular Hepatology. 2019;25:406-15.
21. Li M, Yu B, Yang H, He H, Gao R. Comparative efficacy of non-pharmacological interventions on anxiety, depression, sleep disorder, and quality of life in patients with liver transplantation: a systematic review and network meta-analysis. J Clin Nurs. 2025;34:1993-2010.

Original Article

DOI: 10.4274/haseki.galenos.2026.92408
Med Bull Haseki 2026;64(5):399-406

Sonographic Predictors of Malignancy in BI-RADS 4A Breast Lesions: A Diagnostic Performance Analysis

Gulbanu Guner, Sevdiye Nur Agackaya, Fatma Kulali

University of Health Sciences Türkiye, Umraniye Training and Research Hospital, Clinic of Radiology, Istanbul, Türkiye

Abstract

Aim: Breast Imaging Reporting and Data System (BI-RADS) category 4A lesions present a diagnostic challenge because of their low risk of malignancy and overlapping imaging features. This study aimed to identify sonographic predictors of malignancy in BI-RADS 4A lesions and assess their diagnostic performance.

Methods: This retrospective, single-center observational study included 179 BI-RADS 4A breast lesions from 179 women who were evaluated between January 2024 and January 2025. Ultrasonography (US) images were independently assessed by two radiologists using standardized BI-RADS ultrasound descriptors, and all lesions underwent US-guided core needle biopsy. Multivariate logistic regression was performed to identify independent predictors of malignancy, and receiver operating characteristic (ROC) curves were used to assess diagnostic performance.

Results: Among 179 lesions, 12 (6.7%) were malignant. On univariate analysis, irregular margins (p=0.016), irregular shape (p<0.001), microcalcifications (p=0.018), and parenchymal distortion (p=0.003) were significant. In multivariate analysis, microcalcifications were the only independent predictors (odds ratio: 6.51, 95% confidence interval: 1.66–25.52, p=0.007). ROC analysis showed moderate performance for detecting microcalcifications (area under the curve: 0.66, sensitivity: 42%, specificity: 90%).

Conclusion: Microcalcifications were the only independent sonographic predictor of malignancy in BI-RADS 4A lesions. This finding underscores their importance among conventional ultrasound descriptors and suggests their potential to improve risk stratification and to guide biopsy decisions.

Keywords: Breast neoplasms, ultrasonography, large-core needle biopsy, predictive value of tests

Introduction

Breast cancer remains the most common malignancy among women worldwide and a leading cause of cancer-related mortality (1). Early detection is essential for improving prognosis and survival (2). While mammography is the primary screening modality, its sensitivity decreases in women with dense breast tissue, where ultrasonography (US) plays a crucial complementary role (2,3). The Breast Imaging Reporting and Data System (BI-RADS) provides a standardized lexicon and risk-stratification framework that guide clinical management based on imaging findings (4).

Among BI-RADS categories, category 4 represents a heterogeneous group with a wide malignancy probability (2-95%) and is subdivided into 4A, 4B, and 4C (4,5). BI-RADS 4A lesions, in particular, pose a clinical challenge because their relatively low malignancy rate is often accompanied by biopsy recommendations, leading to a substantial number of benign histopathological outcomes and potentially unnecessary invasive procedures (5,6).

Recent studies suggest that advanced, multimodal ultrasound techniques may improve diagnostic accuracy for BI-RADS 4 and 4A lesions. However, overlap between benign and malignant imaging features remains an important limitation and highlights the need for more reliable predictors of malignancy within BI-RADS 4A lesions (7-10). We hypothesized that specific conventional ultrasonographic descriptors could more accurately predict malignancy within BI-RADS 4A lesions and thereby improve biopsy decision-making.

Corresponding Author: Gulbanu Guner, Specialist, University of Health Sciences Türkiye, Umraniye Training and Research Hospital, Clinic of Radiology, Istanbul, Türkiye
E-mail: gulbanuguner@gmail.com **ORCID:** orcid.org/0000-0001-7825-587X
Received: 16.09.2025 **Accepted:** 01.08.2026 **Publication Date:** 17.11.2026

Cite this article as: Guner G, Agackaya SN, Kulali F. Sonographic predictors of malignancy in BI-RADS 4A breast lesions: a diagnostic performance analysis. Med Bull Haseki. 2026;64(5):399-406

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

This study aimed to evaluate the histopathological outcomes of BI-RADS 4A breast lesions and to identify the ultrasonographic features most strongly associated with malignancy. By doing so, the study aims to contribute to more accurate risk stratification within the BI-RADS 4A category.

Materials and Methods

Compliance with Ethical Standards

The study was approved by the Ethics Committee of University of Health Sciences Türkiye, Umraniye Training and Research Hospital (approval number: 27.02.2025/19, date: 27.02.2025) and was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki. Due to the study's retrospective design, the ethics committee waived the requirement for informed consent.

Study Design and Patient Selection

This retrospective single-center observational study included women with sonographically classified BI-RADS 4A breast lesions who had histopathological results available between January 2024 and January 2025. The inclusion criteria included women with breast lesions categorized as BI-RADS 4A on US who underwent US-guided core needle biopsy and who had complete ultrasonographic images as well as histopathological results. Patients younger than 18 years, those with a history of breast cancer, those without histopathological results, and those with incomplete or insufficient ultrasonographic images were excluded from the study. Of the 345 patients initially assessed for eligibility, 179 patients with BI-RADS 4A lesions were included in the final analysis after applying the inclusion and exclusion criteria. The mean age of the patients was 43±13 years (range, 18-105 years). All lesions were histopathologically confirmed by US-guided core needle biopsy using 14-16-gauge needles (Geotek Medical, Ankara, Türkiye). The patient selection process is illustrated in Figure 1.

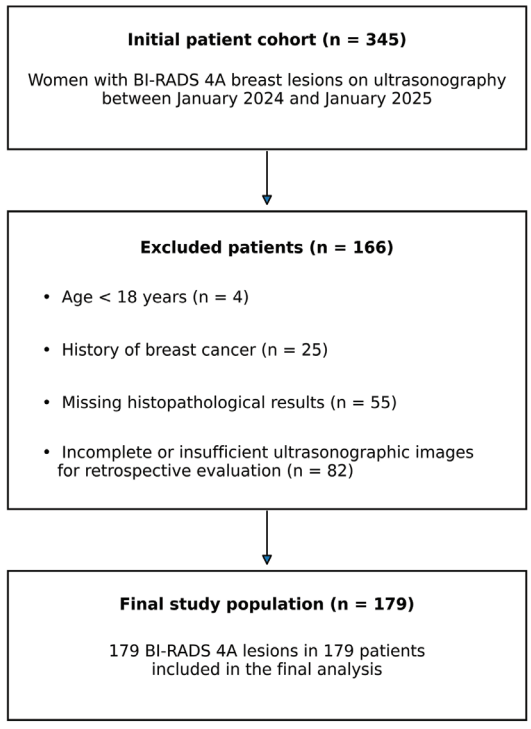


Figure 1. The flowchart about patient selection
BI-RADS: Breast Imaging Reporting and Data System

Ultrasonography Examinations

All examinations were performed using a high-resolution ultrasound system equipped with a linear transducer (central frequency 15 MHz, range 7-18 MHz; Aplio 500; Toshiba Medical Systems, Tokyo, Japan).

Image Evaluation

Ultrasonographic images were independently evaluated by two radiologists with 3 and 8 years' experience in breast imaging, who were blinded to the patients' histopathological results, and consensus was achieved in all cases. In cases of disagreement, the opinion of a third radiologist with 25 years of experience was considered.

The evaluated sonographic parameters were defined according to the BI-RADS ultrasound lexicon (2013) (11) and are summarized in Table 1. These parameters included lesion location, size, margin characteristics (circumscribed, indistinct, microlobulated, ill-defined, or spiculated), composition (solid, cystic, or mixed), echogenicity (hypoechoic, isoechoic, or hyperechoic), posterior acoustic features (none, shadowing, or enhancement), and additional findings such as microcalcifications, parenchymal distortion, and ductal ectasia.

Patient age and BI-RADS classification were also recorded. Although mammography reports were available for some patients, they were not consistently present in all cases and were therefore excluded from the comparative analysis. Histopathological diagnoses were categorized into three groups: benign, high-risk benign, and malignant (Table 2).

Statistical Analysis

Descriptive statistics were expressed as mean ± standard deviation or median (minimum-maximum) for continuous variables and as number and percentage for categorical variables. Normality of data distribution was assessed using the Shapiro-Wilk test.

Continuous variables, including lesion size and age, were compared among groups using the Kruskal-Wallis test. Categorical variables, including margin characteristics, microcalcifications, parenchymal distortion, lesion shape, and posterior acoustic features, were analyzed using the chi-square test or Fisher's exact test when expected cell

Table 1. Distribution of ultrasonographic features of BI-RADS 4A breast lesions according to histopathological subgroup				
Feature	**Benign (n=149) (%)**	**High-risk benign (n=18) (%)**	**Malignant (n=12) (%)**	**p-value**
Margins Circumscribed	118 (79%)	15 (83%)	4 (33%)	0.016*
Irregular	21 (14%)	1 (5.5%)	6 (50%)	
Microlobulated	10 (7%)	2 (11%)	2 (16%)	
Microcalcifications	16 (11%)	2 (11%)	5 (42%)	0.018*
Parenchymal distortion	10 (6.7%)	4 (22%)	6 (50%)	0.003*
Posterior acoustic enhancement	13 (9%)	2 (11%)	0	0.748

Data are presented as number (percentage). p-values were calculated using the chi-square test or Fisher's exact test, as appropriate. *Statistically significant (p<0.05).
BI-RADS: Breast Imaging Reporting and Data System

Table 2. Histopathological distribution of BI-RADS 4A breast lesions	
Pathology	**n (%)**
Benign	**149 (83.3%)**
Fibroadenoma	81 (54.4%)
Fibrocystic changes	51 (34.2%)
Ductal hyperplasia	7 (4.7%)
Mastitis/inflammation	6 (4.0%)
Fat necrosis/fibrosis	4 (2.7%)
High-risk benign	**18 (10.0%)**
Intraductal papilloma	11 (61.1%)
Sclerosing adenosis	5 (27.8%)
Phyllodes tumor	2 (11.1%)
Malignant	**12 (6.7%)**
Invasive breast carcinoma	6 (50.0%)
DCIS	3 (25.0%)
Neuroendocrine carcinoma	1 (8.3%)
Invasive ductal carcinoma	1 (8.3%)
Lobular carcinoma in situ arising from a background of complex fibroadenoma	1 (8.3%)

Data are presented as number (percentage)
BI-RADS: Breast Imaging Reporting and Data System, DCIS: Ductal carcinoma in situ

frequencies were below 5. Histopathological diagnosis was accepted as the reference standard.

Variables found to be statistically significant in univariate analysis (p<0.05) were further evaluated using multivariate logistic regression analysis to identify independent predictors of malignancy, with histopathological outcome defined as a binary dependent variable (malignant vs. non-malignant, including benign and high-risk benign lesions). Odds ratios (ORs) and 95% confidence intervals (CIs) were calculated. Due to potential multicollinearity between shape and margin characteristics, these variables were not included in the same multivariate model; shape was excluded from the final model. Given the limited number of malignant cases, the multivariate logistic regression analysis was considered exploratory.

Receiver operating characteristic (ROC) curve analysis was performed to evaluate the diagnostic performance of sonographic features identified as significant in the univariate analysis. The area under the curve (AUC), along with sensitivity and specificity, was calculated. Post-hoc analysis was not performed, as the primary aim of the study was to identify independent predictors of malignancy, which was addressed using multivariate logistic regression.

All statistical analyses were performed using MedCalc Statistical Software version 12.7.7 (MedCalc Software bvba, Ostend, Belgium).

Results

A total of 179 BI-RADS 4A lesions were included in the study. The mean age of the patients was 43±13 years (range: 18-105 years). The mean age was significantly higher in the high-risk benign and malignant groups than in the benign group (p=0.021).

Histopathological diagnoses were categorized as benign (83.3%), high-risk benign (10.0%), and malignant (6.7%), as summarized in Table 2. Representative ultrasonographic images from each group are shown in Figures 2-4.

No statistically significant difference was observed in lesion size among the benign, high-risk benign, and malignant groups (Kruskal-Wallis test, χ^2=0.48; p=0.787). The mean lesion size was 19.3±11.9 mm in the benign group (range: 3-78 mm), 21.2±17.7 mm in the high-risk benign group (range: 5-63 mm), and 18.8±15.0 mm in the malignant group (range: 4-58 mm).

No statistically significant difference in lesion localization was observed between the groups (chi-square test, p=0.968). Lesions were most frequently located in the upper outer quadrant across all groups, while high-risk benign lesions were relatively predominant in the retroareolar region.

Among the evaluated sonographic features, margin characteristics (p=0.016), microcalcifications (p=0.018), and parenchymal distortion (p=0.003) showed statistically significant differences between the groups (Table 1), whereas posterior acoustic enhancement did

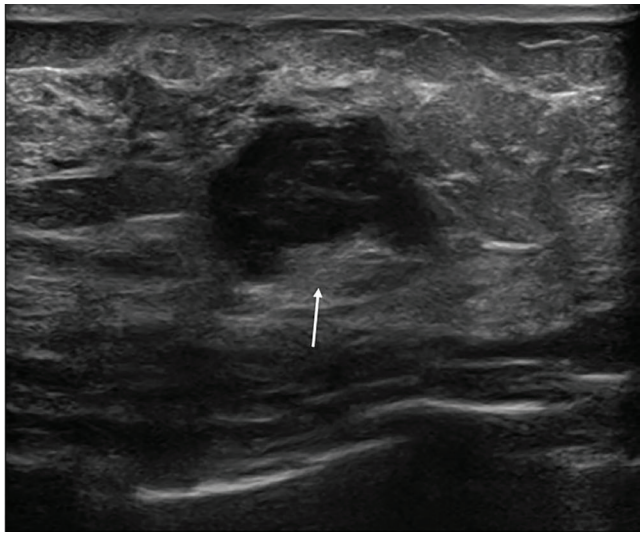


Figure 2. Ultrasound image of a 28-year-old woman demonstrating a hypoechoic solid lesion (indicated by a white arrow) located at the 1 o'clock position of the left breast. The lesion exhibits partially lobulated and partially irregular margins. It was categorized as BI-RADS 4A, and histopathological evaluation confirmed a benign lesion consistent with fibroadenoma
BI-RADS: Breast Imaging Reporting and Data System

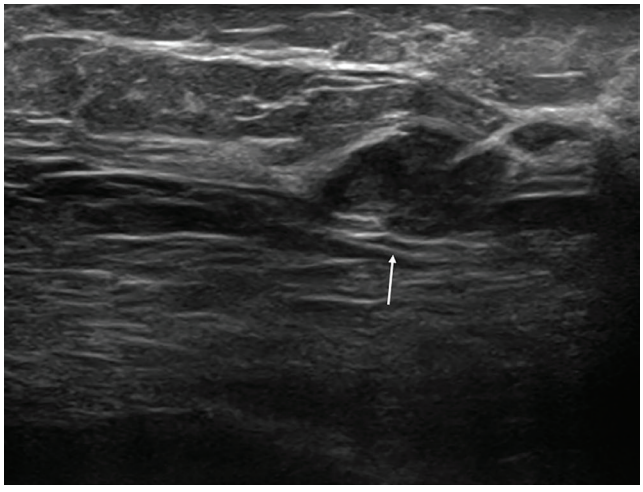


Figure 3. Ultrasound image of a 42-year-old woman demonstrating a lobulated, hypoechoic solid lesion with parallel orientation in the right breast at the 5 o'clock position (indicated by a white arrow). The lesion was categorized as BI-RADS 4A, and histopathological evaluation revealed a high-risk benign lesion consistent with sclerosing adenosis
BI-RADS: Breast Imaging Reporting and Data System

not differ significantly (p=0.748). Margin irregularity, microcalcifications, and parenchymal distortion were more frequently observed in malignant lesions (Table 1).

Lesion shapes differed significantly among the groups (p<0.001). Benign lesions were predominantly oval, whereas irregular shapes were more common in malignant lesions.

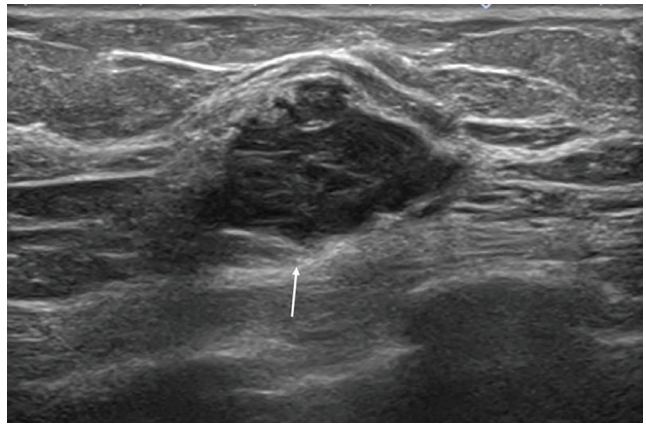


Figure 4. Representative ultrasonographic image of a 55-year-old woman showing a hypoechoic solid lesion with partially indistinct, lobulated, and irregular margins in the left breast at the 2 o'clock position (indicated by a white arrow). The lesion was assessed as BI-RADS 4A on ultrasound, and histopathological examination confirmed a malignant lesion consistent with invasive ductal carcinoma
BI-RADS: Breast Imaging Reporting and Data System

Ductal ectasia did not differ significantly between the groups (p=0.61). Similarly, posterior acoustic shadowing showed no significant difference (p=1.00).

Composition (p=0.928) and echogenicity (p=0.546) also did not differ significantly between the groups.

Variables found to be significant in univariate analysis were further evaluated using multivariate logistic regression. In multivariate analysis, the presence of microcalcifications was identified as an independent predictor of malignancy (OR: 6.51, 95% CI: 1.66-25.52, p=0.007). Although parenchymal distortion (OR: 4.18, 95% CI: 0.57-30.63, p=0.16) and margin characteristics (OR: 2.28, 95% CI: 0.55-9.40, p=0.25) showed increased ORs, neither remained statistically significant in the multivariate model.

Receiver operating characteristic curve analysis was performed to evaluate the diagnostic performance of sonographic features. Microcalcifications demonstrated the highest diagnostic performance (AUC=0.66, 95% CI: 0.50-0.82), followed by lesion shape (AUC=0.64, 95% CI: 0.48-0.80), parenchymal distortion (AUC=0.60, 95% CI: 0.43-0.77), and margin characteristics (AUC=0.56, 95% CI: 0.40-0.72). The overall malignancy rate of BI-RADS 4A lesions in this cohort was 6.7%.

Discussion

This study provides a focused evaluation of sonographic predictors of malignancy in BI-RADS 4A breast lesions, a subgroup characterized by considerable diagnostic uncertainty. Several sonographic features, including irregular margins, irregular shape, parenchymal distortion, and microcalcifications, were significantly associated with

malignancy in univariate analysis. However, in multivariable analysis, microcalcifications remained associated with malignancy; this finding should be interpreted cautiously given the limited number of malignant events. These results highlight the limitations of relying solely on conventional sonographic descriptors and emphasize the necessity of more refined risk-stratification strategies. From a clinical perspective, these findings may contribute to risk stratification within the BI-RADS 4A category.

The reported malignancy rate for BI-RADS 4A lesions ranges from 2% to 10% in the literature (12,13). The malignancy rate observed in our study (6.7%) falls within this range, supporting the representativeness of our cohort. This finding reinforces the notion that although the risk of malignancy is relatively low, BI-RADS 4A lesions remain clinically significant and require careful evaluation.

Advanced age has consistently been identified as a predictor of malignancy in BI-RADS 4A lesions (14). In our study, the higher mean age observed in the malignant and high-risk benign groups further supports this association. From a clinical perspective, these results suggest that patient age may be a complementary parameter for risk stratification.

The association between lesion size and malignancy has been reported inconsistently in previous studies (15,16). However, lesion size alone is not considered a reliable discriminative feature (17). Consistent with these findings, no significant association between lesion size and malignancy was observed, supporting the limited discriminative value of this parameter. Therefore, lesion size should not be considered an independent determinant in clinical decision-making. Our findings further underscore the complexity of risk stratification in BI-RADS 4A lesions.

The ultrasonographic appearance of breast lesions can be highly variable, and both benign and malignant features may coexist within the same lesion (9,18). This heterogeneity contributes to diagnostic uncertainty, particularly in BI-RADS 4A lesions, where overlapping imaging findings may lead to misclassification. In addition, interobserver variability in interpreting BI-RADS descriptors further complicates lesion categorization (19). These factors collectively highlight the inherent limitations of conventional sonographic assessment and underscore the need for more standardized, reproducible evaluation approaches.

Irregular shape is a recognized feature associated with malignancy in breast lesions (20,21). However, it was not included in the multivariate model due to its strong association with margin characteristics, which may have introduced collinearity and affected model stability. Therefore, its independent predictive contribution could not be determined in this study.

Margin characteristics are well-established features associated with malignancy in breast lesions (18). Margin irregularity was more frequently observed in malignant lesions, supporting its role in the initial assessment of malignancy. Although margin characteristics were significantly associated with malignancy in univariate analysis, they did not remain independently significant in multivariate analysis. These results suggest that margin characteristics alone may be insufficient for accurate risk stratification.

Echogenicity has limited discriminative value in distinguishing benign from malignant breast lesions, as substantial overlap exists between groups (22). In the present study, no significant difference in echogenicity was observed, which further supports its limited discriminative value in malignancy prediction.

Cystic morphology is generally associated with benign breast lesions, whereas solid and mixed patterns have limited discriminative value (15,17). In our study, cystic morphology was observed exclusively in benign lesions. In contrast, solid and mixed morphologies were similarly distributed across groups, suggesting a limited role for lesion composition in malignancy differentiation.

Posterior acoustic shadowing is commonly considered a feature suggestive of malignancy; however, it may also be observed in benign lesions, particularly in the presence of fibrosis or calcifications (15). In our study, posterior acoustic features did not show a significant association with malignancy, a finding that may be partly explained by the limited number of malignant cases.

Microcalcifications are well-recognized as a feature associated with malignancy in breast lesions (15,23,24). In our study, microcalcifications were the only parameters that retained independent predictive value in multivariate analysis, highlighting their importance among the evaluated sonographic features. This finding supports the role of microcalcifications as a more reliable feature in BI-RADS 4A lesions compared to other conventional descriptors.

Although microcalcifications were the only feature that remained independently associated with malignancy in the multivariate analysis, they demonstrated moderate diagnostic performance in ROC analysis (AUC=0.66). The lower discriminatory performance of the other sonographic features, with AUC values ranging from 0.56 to 0.64, suggests that microcalcifications, despite their moderate diagnostic performance, may contribute relatively more to risk stratification. From a clinical perspective, the independent association between microcalcifications and malignancy suggests that their presence may serve as an additional indicator of increased malignancy risk in BI-RADS 4A lesions, particularly when accompanied by other

suspicious sonographic findings. Nevertheless, given their moderate diagnostic performance, microcalcifications should not be considered a standalone determinant of biopsy decisions but rather an additional sonographic feature that may support risk assessment within the BI-RADS 4A category.

Recent studies have explored advanced approaches to improve diagnostic accuracy in BI-RADS 4 lesions, including multimodal ultrasound, elastography, contrast-enhanced ultrasound, deep learning, and nomogram-based models (7,8,25,26). While these approaches aim to enhance overall diagnostic performance, they also underscore the limitations of relying solely on individual conventional sonographic features. In this context, our findings indicate that certain features, particularly microcalcifications, may retain a more robust predictive role than other descriptors.

Orientation is generally considered a supportive feature in lesion assessment, with benign lesions more commonly aligned parallel to the skin (27). In our study, however, orientation alone was not a distinctive parameter, suggesting that its contribution to malignancy prediction may be limited when evaluated in isolation.

Parenchymal distortion has been associated with malignancy due to its relation to stromal retraction and infiltrative growth patterns (23). In our study, distortion was more frequently observed in malignant lesions; however, it did not retain statistical significance in multivariate analysis, suggesting that other accompanying features may influence its predictive value.

Previous studies have reported that the most common benign diagnoses in the BI-RADS 4A category are fibroadenoma and fibrocystic changes (17,20). Consistent with these findings, the majority of histopathological diagnoses in our study were fibroadenomas and fibrocystic changes.

Intraductal papilloma was the most common lesion within the high-risk benign group, followed by sclerosing adenosis. These lesions may exhibit imaging features overlapping with malignancy, including irregular margins, microcalcifications, and parenchymal distortion (17,20), potentially contributing to false-positive classification in BI-RADS 4A lesions. Among malignant cases, invasive ductal carcinoma was the most common diagnosis, reflecting its known predominance among breast cancers (28).

Study Limitations

Our study has several limitations. The retrospective single-center design limits the generalizability of the results. In addition, only ultrasonographic findings were evaluated; vascularity and elastographic features of the lesions were not assessed. Breast parenchymal characteristics, including tissue density, were also not evaluated, which may have affected lesion assessment accuracy and contributed to potential misclassification. As the US is an operator-dependent technique with substantial interobserver variability, differences in experience and the learning curve may lead to classification inconsistencies (29). Furthermore, evaluating static images may have led to overlooking dynamic information that real-time US could provide.

Although ultrasonographic images were independently evaluated by two radiologists, only the final consensus assessments were used for analysis. Consequently, interobserver agreement could not be formally assessed using Cohen's kappa or the intraclass correlation coefficient, which constitutes a limitation of the present study.

Another limitation is that benign cases diagnosed by core biopsy were not confirmed by surgical excision or long-term follow-up, potentially resulting in false-negative outcomes. The literature reports a false-negative rate of approximately 3.7% for 14-gauge US-guided core biopsies (30). Because certain sonographic variables were strongly associated, not all features could be included simultaneously in the multivariate model, which may have limited the ability to fully assess the independent predictive contributions of these variables.

Furthermore, the limited number of malignant cases may have affected the stability of the multivariate logistic regression model and increased the risk of overfitting. Therefore, the identified independent predictor should be interpreted with caution and validated in larger, independent cohorts.

Despite these limitations, this study has several strengths. First, all sonographic evaluations were performed independently by radiologists blinded to histopathological results using standardized BI-RADS 2013 ultrasound descriptors, followed by consensus assessment, thereby minimizing potential interpretation bias. Second, histopathological confirmation was available for all included lesions, and no patients were lost to follow-up, ensuring complete diagnostic verification. Third, the detailed categorization of lesions into benign, high-risk benign, and malignant groups reflects real-world clinical practice and provides a clinically relevant framework. Fourth, multivariate analysis was performed after a careful assessment of multicollinearity, allowing the robust identification of independent predictors of malignancy. The observed malignancy rate (6.7%) falls within the reported range for BI-RADS 4A lesions (2-10%), supporting the representativeness of the study sample.

Conclusion

Microcalcifications emerged as the only independent ultrasonographic predictor of malignancy in BI-RADS

4A breast lesions. Although irregular shape, margin characteristics, and parenchymal distortion were significantly associated with malignancy in univariate analysis, only microcalcifications retained independent significance in multivariate analysis. These findings suggest that conventional sonographic descriptors may have limited standalone discriminative value. Emphasizing more reliable imaging features, particularly microcalcifications, may contribute to malignancy risk stratification in BI-RADS 4A lesions.

However, these findings require validation in larger, multicenter prospective studies to confirm their generalizability and clinical applicability.

Ethics

Ethics Committee Approval: The study was approved by the Ethics Committee of Umraniye Training and Research Hospital (approval number: 27.02.2025/19, date: 27.02.2025) and was conducted in accordance with the ethical principles outlined in the Declaration of Helsinki.

Informed Consent: Due to the study's retrospective design, the ethics committee waived the requirement for informed consent.

Footnotes

Authorship Contributions

Concept: G.G., F.K., Design: G.G., S.N.A., Data Collection or Processing: G.G., S.N.A., Analysis or Interpretation: G.G., F.K., Literature Search: G.G., Writing: G.G., F.K.

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

References

1. Siegel RL, Giaquinto AN, Jemal A. Cancer statistics, 2024. CA Cancer J Clin. 2024;74:12-49.
2. Duggan SN, Azharuddin M, Hernández R, et al. Supplemental imaging modalities for breast cancer screening in women with dense breasts: a systematic review with economic considerations. Breast. 2026;85:104668.
3. Gilbert FJ, Payne NR, Allajbeu I, et al. Comparison of supplemental breast cancer imaging techniques-interim results from the BRAID randomised controlled trial. Lancet. 2025;405:1935-44.
4. American College of Radiology. ACR BI-RADS® Atlas—Breast Imaging Reporting and Data System. Reston (VA): American College of Radiology; 2025.
5. Pradeep Kumar D, Chakiath AX, Sugunan A, et al. Rethinking breast imaging reporting and data system (BI-RADS) 4: clinical insights from mammographic subclassification. Cureus. 2025;17:e94379.
6. Castro-Palomares JE, Sanchez-Rodriguez IG, Reyes-Dircio A, et al. Correlation between BI-RADS 4 subcategories and histopathological outcomes in Mexican women with breast lesions: a retrospective study of lesion laterality and cancer incidence. Curr Med Imaging. 2025;21:e15734056407923.
7. Zhu Y, Jiang L, Chen YN, et al. Diagnostic value of real-time ultrasound elastography and contrast-enhanced ultrasound in BI-RADS 4A breast lesions. Gland Surg. 2025;14:488-97.
8. Li C, Wu M, Haiqing H. Application of multimodal ultrasound examination in BI-RADS category 4 breast lesions. Front Med (Lausanne). 2025;12:1596100
9. Zhong Y, Chen YT, Qiu YD, et al. Sonographic glandular tissue component: a potential imaging marker for upgrading BI-RADS 4A breast masses. Acad Radiol. 2025;32:3883-91.
10. Wang J, Ma H, Wang S, et al. Optimizing S-detect classification accuracy for BI-RADS 4 breast nodules using multimodal ultrasound parameters. Quant Imaging Med Surg. 2026;16:159.
11. American College of Radiology. ACR BI-RADS® Atlas: Breast Imaging Reporting and Data System. 5th ed. Reston(VA):American College of Radiology; 2013.
12. Hu Y, Yang Y, Gu R, et al. Does patient age affect the PPV3 of ACR BI-RADS ultrasound categories 4 and 5? Eur Radiol. 2018;28:2492-8.
13. Lazarus E, Mainiero MB, Schepps B, et al. BI-RADS lexicon for US and mammography: interobserver variability and positive predictive value. Radiology. 2006;239:385-91.
14. Ghunaim HA. Effect of patient age and breast parenchymal density on BI-RADS 4 subcategorization. Pak J Med Sci. 2024;40:2356-62.
15. Elverici E, Barca AN, Aktas H, et al. Nonpalpable BI-RADS 4 breast lesions: sonographic findings and pathology correlation. Diagn Interv Radiol. 2015;21:189-94.
16. Bouzghar G, Levenback BJ, Sultan LR, et al. Bayesian probability of malignancy with BI-RADS sonographic features. J Ultrasound Med. 2014;33:641-8.
17. Wiratkapun C, Bunyapaiboonsri W, Wibulpolprasert B, et al. Biopsy rate and positive predictive value for breast cancer in BI-RADS category 4 lesions. J Med Assoc Thai. 2010;93:830-7.
18. Stavros AT, Thickman D, Rapp CL, et al. Solid breast nodules: use of sonography to distinguish between benign and malignant lesions. Radiology. 1995;196:123-34.
19. Mohapatra SK, Mishra A, Sahoo TK, et al. Positive predictive values of BI-RADS 4 lesions and its mammographic features. Indian J Surg Oncol. 2021;12:182-9.
20. Tan YY, Wee SB, Tan MP, et al. Positive predictive value of BI-RADS categorization in an Asian population. Asian J Surg. 2004;27:186-91.
21. Chaitanya INVL, Prabhala S, Srirambhatla A, et al. Comparison of histopathologic findings with BI-RADS score in tru-cut biopsies of breast lesions. Int J Pathol Res Pract. 2020;9:35-41.

22. Hong AS, Rosen EL, Soo MS, et al. BI-RADS for sonography: positive and negative predictive values. AJR Am J Roentgenol. 2005;184:1260-5.

23. Costantini M, Belli P, Lombardi R, et al. Characterization of solid breast masses using the sonographic BI-RADS lexicon. J Ultrasound Med. 2006;25:649-61.

24. Brunetti N, Campi C, Piana M, et al. A radiomic and clinical data-based risk model for malignancy prediction of breast BI-RADS 4A microcalcifications. Clin Breast Cancer. 2025;25:e461-9.

25. Yang Y, Liao T, Lin X, et al. Reducing unnecessary biopsies of BI-RADS 4 lesions using deep learning in mammography. Front Oncol. 2025;15:1543553.

26. Jiang X, Chen C, Yao J, et al. Nomogram for diagnosis of BI-RADS 4 breast nodules based on 3D ultrasound. BMC Med Imaging. 2025;25:48.

27. Shen L, Jiang T, Tang P, et al. Quantitative malignant risk prediction of amorphous calcifications. Quant Imaging Med Surg. 2022;12:2672-83.

28. Schoonjans JM, Brem RF. Fourteen-gauge ultrasound-guided core needle biopsy of breast masses. J Ultrasound Med. 2001;20:967-72.

29. Lee YJ, Choi SY, Kim KS, et al. Observer variability using BI-RADS ultrasound (5th edition). Iran J Radiol. 2016;13:e28281.

30. Crystal P, Koretz M, Shcharynsky S, et al. Accuracy of ultrasound-guided core needle biopsy of breast lesions. J Clin Ultrasound. 2005;33:47-52.

Original Article

DOI: 10.4274/haseki.galenos.2026.59144
Med Bull Haseki 2026;64(5):407-414

Sinoatrial Node Function Remains Unaltered During Transcutaneous Electrical Stimulation: An Electrophysiological Study

Halime Kibar[1], Eser Kalaoglu[1], Ayse Karaoglu[2], Selim Sezikli[1], Tugba Aydin[1], Kemal Sitki Turker[3], Ilhan Karacan[1]

[1]*University of Health Sciences Türkiye, Istanbul Physical Therapy Rehabilitation Training and Research Hospital, Clinic of Physical Therapy and Rehabilitation, Istanbul, Türkiye*

[2]*Istanbul Gelisim University Faculty of Engineering and Architecture, Department of Electrical and Electronics Engineering, Istanbul, Türkiye*

[3]*Istanbul Gelisim University Faculty of Dentistry, Department of Physiology, Istanbul, Türkiye*

Abstract

Aim: Transcutaneous electrical stimulation (TES) is widely used in clinical settings and can produce both direct and indirect effects on the sinoatrial node (SAN). This study aimed to investigate whether TES affects SAN function directly or indirectly via the autonomic nervous system.

Methods: The study involved 24 young, healthy adult volunteers and was conducted in two phases. Experiment 1 examined the direct effects of TES on the SAN, and Experiment 2 examined its indirect effects. In Experiment 1, which used a single-pulse electrical current, participants underwent real and sham TES sessions in random order. The P-QRS response and heart rate variability (HRV) in both time and frequency domains were assessed. In Experiment 2, in which a pulse-train electric current was used, conventional and acupuncture-like transcutaneous electrical nerve stimulation were applied in random order, and HRV was analyzed similarly. In both experiments, the electrical current was applied to the volar aspect of the forearm.

Results: Peristimulus time histograms and waveform averages revealed no heartbeat associated with the electrical pulse. Additionally, time and frequency domain analyses indicated that pulse-train TES had no effect on HRV.

Conclusion: This study is the first to examine the direct effects of TES on SAN action potential and HRV. Results show that low-frequency TES has no direct or indirect impact on heart rhythm or SAN function in healthy individuals and confirm its safety for diagnostic and therapeutic use.

Keywords: Electrical stimulation, transcutaneous electric nerve stimulation, heart rate, tachycardia, bradycardia

Introduction

Transcutaneous electrical stimulation (TES) has been a staple in clinical practice for over a century, used to diagnose and treat various medical conditions. Therapeutic TES, delivered in pulse-train form, is commonly employed in physical therapy, rehabilitation, and professional sports; modalities include neuromuscular electrical stimulation, functional electrical stimulation, and transcutaneous electrical nerve stimulation (TENS) (1,2). Additionally, diagnostic TES, typically in single-pulse form, is frequently used in electromyographic reflex studies to diagnose neuromuscular disorders (3). Although the benefits of TES are well known, its potential side effects on the heart have yet to be convincingly clarified.

Corresponding Author: Eser Kalaoglu, MD, Istanbul Physical Therapy Rehabilitation Training and Research Hospital, Clinic of Physical Therapy and Rehabilitation, Istanbul, Türkiye
E-mail: eserkalaoglu@hotmail.com **ORCID:** orcid.org/0000-0001-8959-6522
Received: 27.04.2025 **Accepted:** 24.03.2026 **Publication Date:** 17.11.2026

Cite this article as: Kibar H, Kalaoglu E, Karaoglu A, et al. Sinoatrial node function remains unaltered during transcutaneous electrical stimulation: an electrophysiological study. Med Bull Haseki. 2026;64(5):407-414

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

The heart's intrinsic pacemaker, the sinoatrial node (SAN), rhythmically generates action potentials propagating through the myocardium, causing the heart to contract. The function of the SAN is modified by the sympathetic and parasympathetic nerve fibers. Electrical current is suggested to indirectly change SAN rhythm through these nerve fibers (4,5). These effects likely dissipate after cessation of electrical stimulation. However, previous observations were conducted several minutes after electrical stimulation rather than during it, resulting in conflicting findings (6-8). In some studies, it has been reported that heart rate variability (HRV) decreases (6,9); in others, it increases or decreases depending on the frequency of the electrical stimulus (8), while in others, it does not change (10).

We hypothesized that electrical current may directly and immediately affect SAN excitability. Owing to the high conduction speed of electrical current, such effects, if present, would be expected to occur within a very short latency window and therefore be temporally locked to the stimulus (i.e., time-locked, synchronous responses). In other words, if TES directly influences action potential generation in the SAN, an electrical stimulus-related P-QRS complex should appear immediately after the stimulus. Such stimulus-locked cardiac responses can be detected using peristimulus time histograms (PSTHs) and waveform-averaging analyses. Unlike previous studies that evaluated heart rate or HRV before and after stimulation—thus primarily reflecting delayed autonomic modulation—the present study specifically tests the immediate, stimulus-synchronous (direct) effect of TES on SAN activity in humans while also examining potential indirect autonomic influences. These findings, by clarifying the cardiac safety profile of TES, may contribute to the safer clinical application of electrotherapy in vulnerable patient populations.

Materials and Methods

Compliance with Ethical Standards

All volunteers provided written informed consent to the experimental procedures, which adhered to the principles outlined in the Declaration of Helsinki and were approved by the Koç University Local Ethics Committee (protocol no: 2017.124.IRB2.038, date: 05.09.2017). Ethical approval was obtained from Koç University because one of the authors, Prof. Dr. Kemal Sitki Turker, was a faculty member at the Koç University at the time the study protocol was approved. Ethical approval was obtained from Koç University because one of the authors, Prof. Dr. Kemal Sitki Turker, was a faculty member at the Koç University at the time the study protocol was approved. The study protocol was registered with ClinicalTrials.gov (NCT05966922, July 13, 2023). This research comprised two phases: Experiment 1 analyzed the impact of a single electrical pulse on HRV, while Experiment 2 explored the effects of TES delivered as pulse trains on HRV.

Study Design and Participants

Healthy volunteers aged 20-45 were recruited for the study; exclusion criteria included pregnancy, alcohol or drug abuse, and use of any medication affecting the cardiovascular system. Twelve male volunteers took part in Experiment 1, and twelve volunteers (one female and eleven males) participated in Experiment 2. The mean ages of participants in Experiments 1 and 2 were 26.3±2.9 and 25.5±6.5 years, respectively. These two experiments were conducted using separate groups to manage the lengthy and complex experimental procedures effectively.

Experimental Procedures

The experimental design, including electrode placement on the forearm and hypothesized neural pathways, is illustrated in Figure 1. Participants rested for 3 minutes in a quiet, temperature-controlled room maintained at 23 °C before each experiment. They were asked to abstain from strenuous exercise and caffeine consumption for at least 2 hours. Additionally, participants were advised to refrain from activities such as talking, coughing, and deep breathing during measurements to minimize potential influences on the results.

During electrical stimulation, participants were seated with their elbows flexed at 90 degrees and their hands supinated. Self-adhesive electrodes (4 x 4 cm) were positioned on the flexor aspect of the right forearm to deliver the electrical current (Figure 1). Heart rate variability changes during rest and stimulation were monitored with a three-lead electrocardiogram (ECG).

Experiment 1 (Testing a Single Electrical Pulse on Heart Rate Variability)

Participants underwent both TES and Sham TES sessions in random order. Each session lasted 600 seconds, with a 5-minute rest interval between sessions. Sham TES was performed to observe random variations in the occurrence of the P-QRS complex over time.

Throughout the experiment, the software (Spike2) used the peak of the R wave within the P-QRS complex observed in the ECG channel as a trigger. This triggering allowed the delivery of an electrical stimulus within a 20-250 ms interval before the next P-wave. Trial recordings substantiated that the electrical stimulation consistently occurred 20-250 ms before the onset of the P-wave.

A single monophasic square electrical pulse (i.e., unidirectional pulsed current, also termed interrupted direct current/galvanic) was administered at random intervals of 1 to 4 seconds using a constant-current stimulator (model DS7A, Digitimer Ltd, Hertfordshire, UK) with a pulse width of 200 microseconds. Participants received the highest current intensity they could tolerate,

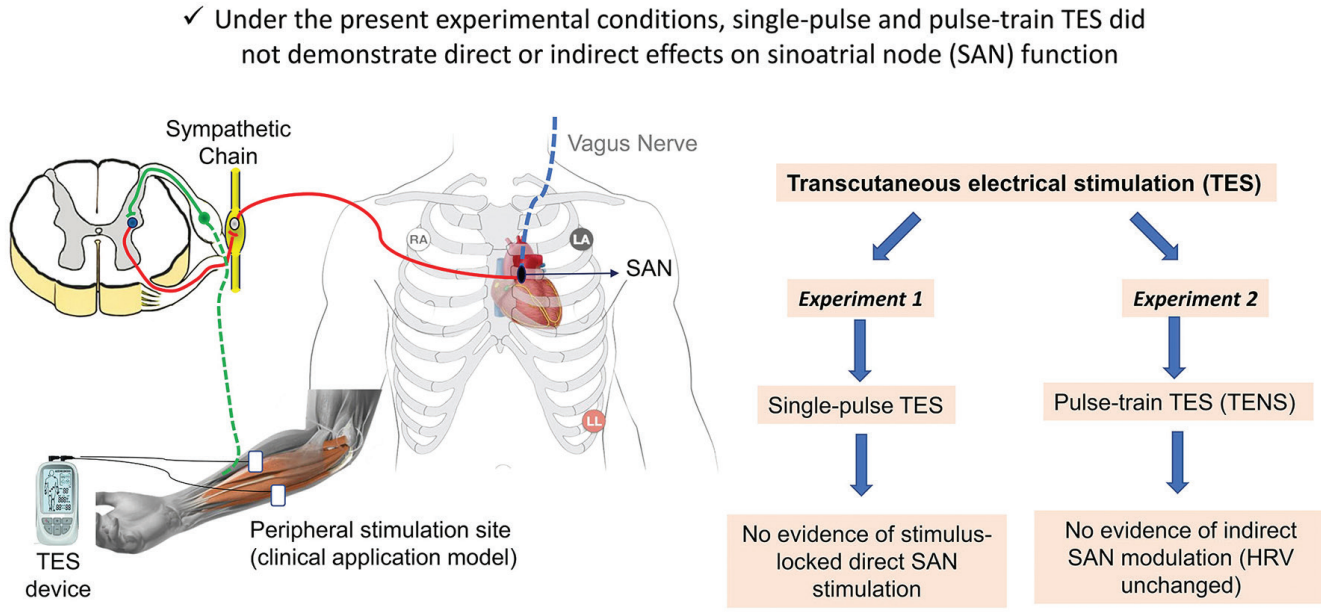


Figure 1. Experimental design and investigated neural pathways. TES was applied to the forearm to investigate potential direct and indirect effects on SAN function. The diagram illustrates the hypothesized neural pathways -the sympathetic chain (red) and the vagus nerve (blue)- and the two-phase experimental design. Experiment 1 tested for direct electrical effects using single-pulse stimulation, while Experiment 2 assessed indirect autonomic effects using pulse-train stimulation

with a mean intensity of 48.1±13.3 milliamperes (mA), ranging from 30 to 70 mA. During the sham stimulation session, the current intensity was set to zero mA.

Experiment 2 (Testing Transcutaneous Electrical Nerve Stimulation on Heart Rate Variability)

In this experiment, conventional and acupuncture-like TENS, which are commonly used in clinical practice, were employed as pulse-train TENS modalities. Initially, a resting ECG was recorded for 800 seconds. Subsequently, one session of conventional TENS and one session of acupuncture-like TENS were applied in random order. Each TENS session lasted 800 seconds, with a 5-minute rest interval between sessions.

The conventional mode is characterized by a high-frequency (HF) current (100 Hz, 200 μs pulse width, 4-second on period, and 10-second off period). The stimulus intensity was adjusted to a level at which the subject experienced minimal paresthesia without muscle contractions. If the tingling sensation faded, the current intensity was increased until the sensation was perceived again. The acupuncture-like mode is characterized by a low frequency (LF) current (10 Hz, 200 μs pulse width, 4 seconds on period, 10 seconds off period). The stimulus intensity was adjusted to a level that allowed for comfortable contractions of the wrist flexors. TENS was administered using the CefarRehab® X2 device (DJO France SAS, Mouguerre, France).

Three-lead ECG Recording

The participant's ECG data were continuously recorded using non-invasive sensors (self-adhesive pre-gelled Ag/AgCl electrodes, Redline® Istanbul, Turkiye) with standard three-lead ECG chest placement. Specifically, the active electrode was placed at the right second intercostal space and the reference electrode at the left second intercostal space, both along the midclavicular line. Another active electrode was situated on the anterior chest wall at the left lower rib cage. Before electrode placement, the skin was shaved, lightly abraded, cleaned with alcohol, and moistened with ECG gel. Skin resistance was measured to ensure an optimal connection (less than 10 kiloohms) (11).

ECG signals were recorded using a data acquisition system (CED 1902 Quad System MKIII amplifier and CED 3601 Power 1401 MKII digital-to-analog converter, Cambridge Electronic Designs, Cambridge, UK). The ECG potentials were amplified 100 times and recorded at a sampling frequency of 2 kHz.

Outcome Variables and Recording Analysis

The primary outcome variables differed between the two experiments because they addressed distinct physiological mechanisms and time scales of potential TES effects on the SAN. In Experiment 1, the main outcome was the electrical stimulus-related QRS response,

assessed using PSTHs and waveform-averaging analyses to determine whether TES could evoke an immediate, stimulus-locked cardiac response indicative of a direct effect on SAN excitability. Beat-to-beat RR interval changes (i.e., HRV) were analyzed to explore short-term variability accompanying single-pulse stimulation.

In contrast, Experiment 2 was designed to evaluate the possible indirect autonomic influence of pulse-train TES on SAN function over time; therefore, the primary outcome was HRV derived from RR interval fluctuations in the time and frequency domains. Thus, while both experiments involved ECG-derived measures, Experiment 1 focused on event-related, stimulus-synchronous QRS responses, whereas Experiment 2 examined slower, autonomically mediated modulation of cardiac rhythm, reflected in HRV.

The Electrical Stimulus-Related QRS Response

If TES affects cardiac rhythm, it would be expected to advance or delay the timing of the QRS response. The possible direct effect of an electrical stimulus on the timing of the QRS response was investigated in the time domain using the PSTH and waveform averaging (11). If a typical P-QRS complex, similar to the one just before stimulation, was evident in the waveform-averaged graphs after electrical stimulation, or if a very narrowly distributed bell curve appeared in the PSTH graphs compared with sham, the observation was interpreted as an electrical current-related P-QRS complex.

Heart Rate Variability

The time-domain analysis assessed beat-to-beat interval variability (i.e., R-R interval differences). This analysis included the standard deviation (SD) of the R-R intervals (SDRR), the coefficient of variation of the R-R intervals (CVRR), and the square root of the mean of the squared differences between adjacent R-R intervals (RMSSD). Prolongation of the R-R interval or an increase in R-R interval variability (SDRR, CVRR, and RMSSD) was interpreted as an increase in parasympathetic tone, reflecting greater parasympathetic (vagal) modulation of cardiac rhythm, which in healthy individuals is generally associated with more adaptive autonomic regulation (12).

Frequency-domain HRV was analyzed in accordance with the guidelines of the Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. The fast fourier transform was used to calculate the LF (0.04-0.15 Hz) and HF (0.15-0.4 Hz) components. Low-frequency and HF components were considered to reflect sympathetic and parasympathetic activity, respectively. The sympathovagal balance was assessed by the LF/HF ratio, where ratios greater than 1 indicated sympathetic dominance (sympatheticotonia), ratios less than 1 indicated parasympathetic dominance (vagotonia), and a ratio of 1 indicated sympathovagal balance (11). ECG data were analyzed using the Spike2 7.20 software package (CED, UK).

Statistical Analysis

An a priori power analysis was conducted using G*Power (version 3.1.9.7) for a paired-samples t-test. The sample size was calculated based on the study by Ngai and Jones (13), which reported an effect size of 1.25 for the LF variable. This effect size was considered appropriate because the present study was designed to detect clinically meaningful alterations in HRV rather than subtle physiological fluctuations. Assuming a power of 95% and an alpha error of 5%, the required sample size was 11 subjects per experiment.

The normality of the data distribution was assessed using the Shapiro-Wilk test. Data were summarized as arithmetic means and SDs. Paired t-tests or repeated-measures analysis of variance (ANOVA) were used to compare the mean values across experimental sessions. Post-hoc pairwise comparisons with Bonferroni correction were performed only when the overall repeated-measures ANOVA yielded a statistically significant effect. Categorical dependent variables were compared using Cochran's Q test for related samples. Statistical significance was set at $p<0.05$. Data management and analysis were conducted using PASW Statistics for Windows, version 18.0 (SPSS Inc., Armonk, NY).

Results

Data analysis from Experiment 1, using PSTHs and waveform averages, indicated no heartbeat related to the electrical pulse. Waveform-averaging analysis revealed that a typical P-QRS complex did not occur in any of the participants after electrical stimulation. No difference was detected in PSTH histogram patterns between electrical stimulation and sham stimulation (Figure 2). Furthermore, there were no statistically significant differences between electrical and sham stimulation for HRV variables in both the time and frequency domains (Table 1).

Analysis of Experiment 2 data revealed no statistically significant differences in time-domain and frequency-domain HRV variables among the resting state, conventional TENS, and acupuncture-like TENS (Table 2). Although LF, HF, and the LF/HF ratio showed borderline p-values in the repeated-measures ANOVA, no post-hoc pairwise comparisons were conducted because the overall ANOVA did not reach statistical significance.

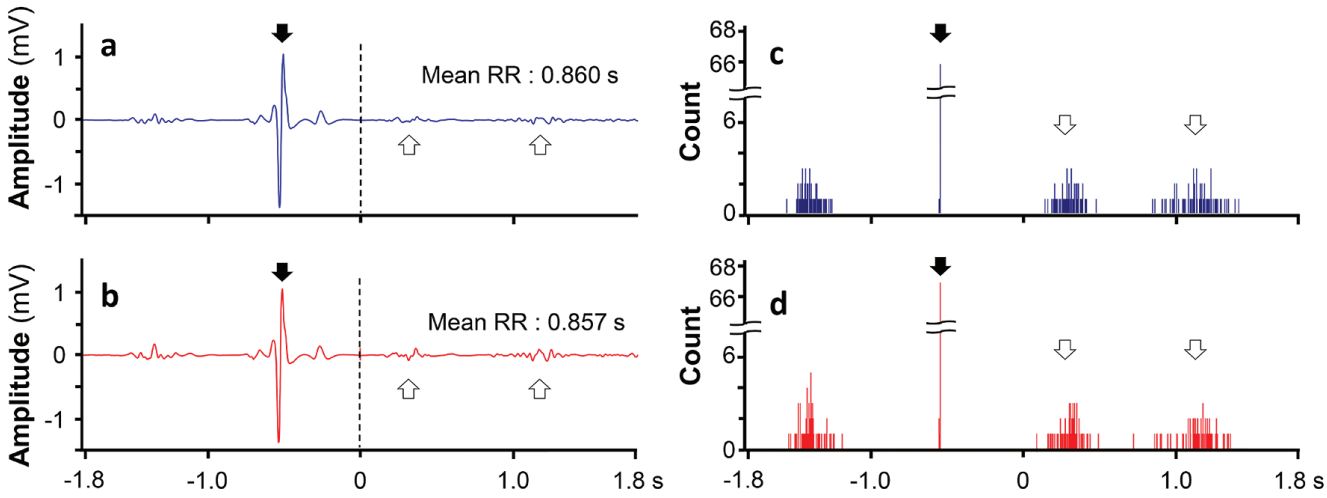


Figure 2. Waveform average plots (**a, b**) and peristimulus time histograms (PSTH; **c, d**) from a representative subject. The upper panels (**a, c**) correspond to sham stimulation, whereas the lower panels (**b, d**) correspond to real stimulation. Vertical dashed lines indicate the timing of electrical pulse delivery. P-QRS complexes immediately preceding stimulation are marked with solid arrows, whereas those occurring after stimulation are indicated with open arrows. No stimulus-locked P-QRS complex was observed following electrical stimulation, and the PSTH distributions were comparable between the sham and real stimulation conditions

Table 1. Heart rate variability in the time and frequency domain in Experiment 1				
		TES	**Sham TES**	**p-value**
Average RR (ms)		762.6±70.4	749.8±70.6	0.095
SDRR (ms)		47.2±12.0	50.6±12.3	0.081
CVRR (%)		6.3±1.5	6.7±1.7	0.104
RMSSD (ms)		28.5±9.8	28.2±8.3	0.857
Low frequency (nu)		75.5±9.2	78.1±7.8	0.353
High frequency (nu)		25.0±8.5	22.0±7.5	0.225
Low/High frequency ratio		3.5±1.6	4.0±1.6	0.249
Low/High frequency (N)	**>1**	12	12	NS
	<1	0	0	

Data are presented as mean ± standard deviation. p-values were calculated using a paired-samples t-test. Statistical significance was set at p<0.05
NS: Non-significant, nu: Normalized units, N: Number of subjects, ms: Millisecond, SDRR: Standard deviation of the R-R intervals, CVRR: Coefficient of variation of the R-R intervals, RMSSD: Square root of the mean of the sum of squares of differences between adjacent R-R intervals, TES: Transcutaneous electrical stimulation

Table 2. Effects of TENS on the heart rate variability in the time and frequency domain						
		Rest	**Acupuncture like TENS**	**Conventional TENS**	**p-value**	**Partial eta squared**
Average RR (ms)		740.3±67.5	741.0±63.0	755.3±66.8	0.161*	0.161
SDRR (ms)		47.8±17.3	46.9±7.5	46.7±11.8	0.873*	0.006
CVRR (%)		6.4±2.1	6.3±0.8	6.2±1.5	0.845*	0.015
RMSSD (ms)		27.4±11.4	27.8±8.1	29.3±10.5	0.444*	0.061
Low frequency (nu)		64.7±12.4	72.3±9.8	72.9±9.2	0.054*	0.253
High frequency (nu)		34.9±11.1	27.7±9.5	27.0±8.5	0.057*	0.263
Low/High frequency ratio		2.2±1.1	3.1±1.7	3.1±1.4	0.057*	0.249
Low/High frequency (N)	**>1**	11	12	12	0.368**	
	<1	1	0	0		

*Repeated measures of variance analysis, **Related-samples Cochran's Q test
TENS: Transcutaneous electrical nerve stimulation, SDRR: Standard deviation of the R-R intervals, CVRR: Coefficient of variation of the R-R intervals, RMSSD: Square root of the mean of the sum of squares of differences between adjacent R-R intervals

Discussion

This study, which included young, healthy adult volunteers, yielded two original findings: PSTH and waveform-average analyses of ECG data recorded during the TES session showed that the electrical current had no direct effect on SAN function. Secondly, HRV analysis of ECG data recorded during the TES session showed that the electrical current had no indirect effect on SAN function via the autonomic nervous system (ANS).

Importantly, this methodological distinction has critical implications for interpreting the existing literature. Studies relying on pre-post HRV comparisons may capture delayed or non-specific autonomic adjustments rather than the immediate physiological response to stimulation (6-10). By contrast, the present study isolates both stimulus-synchronous electrophysiological responses and real-time autonomic modulation during TES. Therefore, the absence of both a stimulus-locked P-QRS response and a concurrent HRV change provides converging evidence that TES, under the present conditions, does not exert a functionally meaningful influence on SAN activity. Importantly, this should be interpreted as evidence of absence within the temporal resolution and sensitivity of the applied methods, rather than a definitive exclusion of all possible effects under different stimulation conditions.

To assess the effect of TES on sinus node function via the ANS, we did not apply stimulation to sympathetic ganglia, such as the stellate ganglion, in this study. Instead, TES was delivered to the forearm, consistent with its common peripheral application in clinical settings. This approach was chosen to reflect typical TES usage and to avoid direct stimulation of cardiac-related autonomic structures.

To directly stimulate the SAN, an electrical pulse must reach it before the onset of the atrial action potential's absolute refractory period. In Experiment 1, an electrical pulse locked to the R-wave was delivered to ensure this. Specifically, the R-wave was used as the trigger, and a single electrical pulse was delivered within an interval of 20-250 ms before the subsequent P-wave. ECG recordings obtained during TES were analyzed using PSTH and waveform-averaging techniques to detect the stimulus-related P-QRS response; this response was then compared with responses obtained during sham stimulation.

The SAN function is modulated by both sympathetic and parasympathetic nerve fibers (5). Experiment 2 tested the indirect effect of TES on the SAN via the ANS. HRV, a non-invasive measure of ANS activity, was used as an objective index for assessment (14). While previous studies typically analyzed HRV before and after TES, this study analyzed HRV from ECG recordings obtained during both rest and TENS sessions, which distinguishes it from prior research. This methodological distinction is critical, as post-stimulation HRV measurements may reflect residual autonomic responses, relaxation effects, or postural changes rather than the direct neuromodulatory effects of TES.

During Experiment 1, no stimulus-locked, pulse-related P-QRS response was observed, indicating that the electrical current did not directly excite the SAN. While previous TES studies have primarily evaluated delayed changes in heart rate or HRV reflecting autonomic modulation, the present study specifically examined the immediate, stimulus-synchronous electrophysiological response of the SAN. To our knowledge, this time-locked approach to testing a direct effect of TES on the SAN has not previously been reported in humans.

The alpha motor neuron regulates the functions of skeletal muscle. In electrophysiological studies, the alpha motor neuron or its axon is stimulated with TES, and a skeletal muscle response associated with an electric pulse is observed in electromyography (15). Similarly, if the SAN can be stimulated with TES, a P-QRS complex associated with an electrical pulse is expected to occur on the ECG. PSTH and waveform averaging are commonly used methods to show stimulus-related responses (15). In Experiment 1, ECG recordings taken during TES did not reveal a P-QRS response associated with an electrical pulse. Figure 2 illustrates that the histogram and P-QRS patterns generated using the PSTH and waveform-average graphs of the ECG recordings taken during real and sham stimulations are remarkably similar. These analyses also revealed that a typical P-QRS wave was not observed after TES. These findings indicate that TES does not have a direct effect on the SAN.

It has been observed that TES activates the ANS, as evidenced by changes in systolic and diastolic blood pressure and skin temperature, as well as sympathetic skin response measurements (16-19). However, not all systems regulated by the ANS respond to TES simultaneously. For instance, Labrunée et al. (20) reported that both TENS and neuromuscular electrical stimulation were able to reduce muscle sympathetic nerve activation compared to sham stimulation, but no changes in blood pressure, heart rate, or respiratory parameters were observed after stimulation. In the study by Donadio et al. (17), which involved 19 healthy participants, TES stimulation of the left middle finger decreased electrical skin resistance. However, the inhibitory effect on muscle sympathetic nerve activity was not observed in all subjects (17). In summary, even if TES activates the ANS, this effect may not be reflected in heart rate or rhythm. The absence of HRV changes in the present study may be explained by the fact that measurements were obtained during stimulation rather than after it, thereby eliminating the confounding influence of post-stimulation autonomic rebound (7).

In neurophysiological practice, responses of excitable tissues to electrical stimulation are tested by administering current. In the second experiment, HRV analyses in both the time and frequency domains revealed that neither conventional nor acupuncture-like TENS affected HRV compared with resting recordings. In the literature, studies examining the effect of TES on HRV employ a pre- and post-stimulation design with HRV evaluations and report conflicting results. In studies reporting changes in HRV after TES application, these changes may be affected by various factors, such as relaxation at the end of TES application, removal of electrodes from the skin, and changes in the subject's posture. From a clinical safety perspective, the absence of any TES-induced HRV changes during stimulation suggests that standard therapeutic TES parameters do not acutely perturb cardiac autonomic regulation in healthy adults, which has important implications for the safe use of electrotherapy in rehabilitation settings (2).

Study Limitations

This study has several limitations that should be acknowledged. First, the study only included young, healthy participants, which limits the generalizability of the findings to clinical populations. The effects of TES on individuals with heart conditions, particularly those with rhythm disorders or autonomic dysfunction, and on elderly individuals remain to be investigated. Second, because electrode placement was restricted to the forearm, the electric field distribution and the actual field strength reaching deeper thoracic structures, including the SAN, could not be directly measured. Third, only two TENS modalities were tested; other stimulation parameters such as different frequencies, intensities, or electrode placements may yield different results. Despite these limitations, this study has notable strengths. It is the first study to examine the direct effect of TES on SAN excitability during stimulation, using PSTH and waveform averaging analyses, a methodological approach previously applied to skeletal muscle but never to cardiac tissue. Furthermore, the crossover design, sham-controlled protocol, and simultaneous ECG recording during stimulation provide a rigorous framework that addresses the key methodological gap identified in prior literature.

Conclusion

Theoretically, side effects rather than positive effects may occur during TES application. Potential issues, such as arrhythmia (dysrhythmia), as well as more serious complications like syncope, could raise safety concerns. This study is the first to examine the SAN action potential generation and to assess HRV during TES administration. The results indicate that low-frequency (<200 Hz) TES, commonly used for diagnostic and therapeutic purposes, does not disrupt heart rhythm or SAN function, either directly or indirectly. Electrical stimulation, as used in the current study, is safe. However, to confirm the cardiac safety of TES, particularly in older adults and individuals with cardiac arrhythmias, further testing during its application is necessary.

Ethics

Ethics Committee Approval: Approved by the Koç University Local Ethics Committee (protocol no: 2017.124.IRB2.038, date: 05.09.2017). Ethical approval was obtained from Koç University because one of the authors, Prof. Dr. Kemal Sitki Turker, was a faculty member at the Koç University at the time the study protocol was approved.

Informed Consent: All volunteers provided written informed consent for the experimental procedures.

Acknowledgments

The authors would like to thank all the participants who contributed to this study.

Footnotes

Authorship Contributions

Concept: H.K., E.K., A.K., S.S., T.A., K.S.T., I.K., Design: H.K., E.K., A.K., S.S., T.A., K.S.T., I.K., Data Collection or Processing: H.K., K.S.T., I.K., Analysis or Interpretation: E.K., A.K., Literature Search: E.K., S.S., Writing: H.K., E.K., A.K., S.S., T.A., K.S.T., I.K.

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

References

1. Bellew JW, Michlovitz SL, Nolan Jr TP. Michlovitz's modalities for therapeutic interventions. Philadelphia, FA Davis Company; 2016.
2. Ono S, Kato M, Seko H, et al. Randomised crossover trial of home-based neuromuscular electrical stimulation therapy as an adjunct to cardiac rehabilitation in frail older adult patients with chronic heart failure. Circ Rep. 2025;7:231-8.
3. Rampello L, Rampello L, Arcidiacono A, Patti F. A waves in electroneurography: differential diagnosis with other late responses. Neurol Sci. 2020;41:3537-45.
4. Campos MO, Nóbrega ACL, Miranda SM, et al. Transcutaneous electrical nerve stimulation attenuates cardiac sympathetic drive in heart failure: a 123MIBG myocardial scintigraphy randomized controlled trial. Am J Physiol Heart Circ Physiol. 2019;317:H226-33.
5. MacDonald EA, Rose RA, Quinn TA. Neurohumoral control of sinoatrial node activity and heart rate: insight from experimental models and findings from humans. Front Physiol. 2020;11:170.

6. Cho SH. Frequency and intensity of electrical stimulation of human sympathetic ganglia affect heart rate variability and pain threshold. App Sci. 2019;9:4490.
7. do Amaral Sartori S, Stein C, Coronel CC, Macagnan FE, Plentz RDM. Effects of transcutaneous electrical nerve stimulation in autonomic nervous system of hypertensive patients: a randomized controlled trial. Curr Hypertens Rev. 2018;14:66-71.
8. Stein C, Dal Lago P, Ferreira JB, Casali KR, Plentz RD. Transcutaneous electrical nerve stimulation at different frequencies on heart rate variability in healthy subjects. Auton Neurosci. 2011;165:205-8.
9. Vieira PJ, Ribeiro JP, Cipriano G Jr, et al. Effect of transcutaneous electrical nerve stimulation on muscle metaboreflex in healthy young and older subjects. Eur J Appl Physiol. 2012;112:1327-34.
10. Kang JH, Hyong IH. The influence of neuromuscular electrical stimulation on the heart rate variability in healthy subjects. J Phys Ther Sci. 2014;26:633-5.
11. Türker KS, Powers RK. Effects of large excitatory and inhibitory inputs on motoneuron discharge rate and probability. J Neurophysiol. 1999;82:829-40.
12. Heart rate variability: standards of measurement, physiological interpretation and clinical use. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology. Circulation. 1996;93:1043-65.
13. Ngai SP, Jones AY. Changes in skin impedance and heart rate variability with application of Acu-TENS to BL 13 (Feishu). J Altern Complement Med. 2013;19:558-63.
14. Cipriano G Jr, Neder JA, Umpierre D, et al. Sympathetic ganglion transcutaneous electrical nerve stimulation after coronary artery bypass graft surgery improves femoral blood flow and exercise tolerance. J Appl Physiol (1985). 2014;117:633-8.
15. Yildiz N, Cecen S, Sancar N, Karacan I, Knikou M, Türker KS. Postsynaptic potentials of soleus motor neurons produced by transspinal stimulation: a human single-motor unit study. J Neurophysiol. 2024;131:1101-11.
16. Moreira BR, Duque AP, Massolar CS, et al. Transcutaneous electrical stimulation of PC5 and PC6 acupoints modulates autonomic balance in heart transplant patients: a pilot study. J Acupunct Meridian Stud. 2019;12:84-9.
17. Donadio V, Kallio M, Karlsson T, Nordin M, Wallin BG. Inhibition of human muscle sympathetic activity by sensory stimulation. J Physiol. 2002;544:285-92.
18. Ebersold MJ, Laws ER Jr, Albers JW. Measurements of autonomic function before, during, and after transcutaneous stimulation in patients with chronic pain and in control subjects. Mayo Clin Proc. 1977;52:228-32.
19. Wong RA, Jette DU. Changes in sympathetic tone associated with different forms of transcutaneous electrical nerve stimulation in healthy subjects. Phys Ther. 1984;64:478-82.
20. Labrunée M, Despas F, Marque P, et al. Acute electromyostimulation decreases muscle sympathetic nerve activity in patients with advanced chronic heart failure (EMSICA Study). PLoS One. 2013;8:e79438.

Original Article

DOI: 10.4274/haseki.galenos.2026.42714
Med Bull Haseki 2026;64(5):415-425

Temporal Muscle Thickness and Long-term Survival after Surgical Treatment of Chronic Subdural Hematoma: A 10-year Single-center Retrospective Cohort Study

Bilal Bahadir Akbulut[1], Elif Ezgi Cenberlitas[2], Mustafa Serdar Boluk[2], Huseyin Biceroglu[2], Erkin Ozgiray[2], Taskin Yurtseven[2]

[1]*Izmir Provincial Directorate of Health, Bornova Turkan Ozilhan State Hospital, Clinic of Neurosurgery, Izmir, Türkiye*

[2]*Ege University Faculty of Medicine, Department of Neurosurgery, Izmir, Türkiye*

Abstract

Aim: Temporal muscle thickness (TMT) has shown inconsistent prognostic utility in chronic subdural hematoma (cSDH). This study evaluated whether TMT independently predicted survival among surgically treated patients with cSDH while controlling for established clinical markers.

Methods: This is a 10-year single-center retrospective cohort study (2013-2023) that analyzed 105 patients aged ≥18 years who underwent surgical cSDH drainage at a tertiary center. Two independent neurosurgeons measured TMT on preoperative computed tomography scans, demonstrating excellent interrater reliability (intraclass correlation coefficients >0.9). TMT was adjusted for age and sex using linear regression in our cohort. Multivariate Cox regression identified independent predictors of mortality while controlling for traditional clinical markers.

Results: The mean age was 69.6±14.4 years with a mean TMT of 5.6±1.6 mm. Survival rates were 96.2% at 3 months and 63.7% at 60 months. In multivariate analysis, independent predictors of mortality included age [hazard ratio (HR)=1.06, p<0.005], albumin levels (HR=0.92, p=0.05; p=0.03 in the model without TMT), keyhole surgical technique (HR=0.18, p<0.005), living status (HR=2.58, p=0.04), and bilateral hematomas (HR=3.23, p=0.02). Temporal muscle thickness was not an independent predictor (HR=0.94, p=0.61).

Conclusion: TMT does not independently predict cSDH survival when traditional clinical markers are considered. Age, albumin levels, surgical technique, living status, and hematoma laterality remain stronger prognostic determinants than TMT alone.

Keywords: Chronic subdural hematoma, survival analysis, temporal muscle, prognosis

Introduction

Chronic subdural hematoma (cSDH) is among the most frequently encountered neurosurgical conditions worldwide. A recent systematic review of the global literature, covering 18 studies across 12 countries and five continents, reported clinical and surgical incidence rates ranging from 3.39 to 39.1 per 100,000 person-years, with incidence rising up to threefold among patients over 80 years of age; notably, no reliable prevalence data currently exist (1). Its incidence continues to increase alongside global population aging, placing increasing demands on neurosurgical services (2). Historically, cSDH has been associated with significant mortality rates, ranging from 9 to 33%, depending on patient characteristics and treatment modalities (3). Several established predictors of poor outcomes in cSDH have been identified, including advanced age, comorbidity burden (4), anticoagulant use, initial neurological status, and radiographic features such

Corresponding Author: Bilal Bahadir Akbulut, MD, Izmir Provincial Directorate of Health, Bornova Turkan Ozilhan State Hospital, Clinic of Neurosurgery, Izmir, Türkiye
E-mail: b.bahadirakbulut@gmail.com **ORCID:** orcid.org/0000-0002-7983-5056
Received: 21.08.2025 **Accepted:** 23.08.2026 **Publication Date:** 17.11.2026

Cite this article as: Akbulut BB, Cenberlitas EE, Boluk MS, Biceroglu H, Ozgiray E, Yurtseven T. Temporal muscle thickness and long-term survival after surgical treatment of chronic subdural hematoma: a 10-year single-center retrospective cohort study. Med Bull Haseki. 2026;64(5):415-425

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

as hematoma volume and midline shift (3). Recently, there has been more interest in using measures of patient frailty as possible prognostic indicators. For example, temporal muscle thickness (TMT) for sarcopenia has become a crucial marker in neurosurgical populations (5-11). Its availability on standard cranial computed tomography (CT) imaging, without extra radiation exposure or special protocols, makes it a viable choice for opportunistic prognostic assessment. The recent multidisciplinary guideline for cSDH care emphasizes frailty-informed risk stratification throughout the patient pathway, yet recommends no imaging-based sarcopenia marker for prognostication (12).

However, evidence regarding TMT's prognostic utility in cSDH remains inconsistent. Korhonen et al. (8) reported that reduced TMT independently predicted shorter survival in patients with cSDH, proposing sex-specific cut-off values for mortality risk stratification. Dubinski et al. (7) associated reduced TMT with poorer functional outcomes but did not address long-term survival. Du et al. (6) raised important methodological concerns regarding these findings, questioning whether TMT alone provides sufficient frailty assessment without incorporating additional body composition metrics or controlling for surgical variables. These discrepancies raise important questions about whether TMT provides independent prognostic information beyond established clinical markers. We hypothesized that TMT would not independently predict mortality among surgically treated cSDH patients, even after adequate adjustment for traditional clinical and radiological variables.

In this study, we present our 10-year retrospective review evaluating TMT as a prognostic marker in patients undergoing surgical drainage for cSDH. We sought to determine whether TMT independently predicts mortality, even after adjustment for established clinical parameters, including previously unexamined influences of surgical technique and hematoma laterality. We assessed whether post-operative changes in TMT provided additional prognostic information beyond baseline measurements. By addressing these questions, we aim to refine our understanding of TMT's role in cSDH outcomes and support more targeted patient risk stratification.

Materials and Methods

Compliance with Ethical Standards

Ethical approval was obtained from the Ege University Ethics Committee (approval number: 24-3T/34, date: 07.03.2024). The patients' consent to participate was obtained using the ethics committee-approved form. This study was conducted in accordance with the ethical standards outlined in the 1964 Declaration of Helsinki and its subsequent amendments.

Study Design

This retrospective cohort study evaluated patients aged 18 years or older who underwent surgical drainage for cSDH at the Department of Neurosurgery, Ege University Faculty of Medicine, between January 2013 and January 2023. This time frame was chosen to ensure a minimum follow-up period of two years and to coincide with the full implementation of our electronic recording system in 2013. The patient selection flowchart is shown in Figure 1.

Chronic subdural hematoma was diagnosed based on clinical presentation with neurological symptoms attributable to the subdural collection, combined with CT evidence of subdural fluid collection demonstrating >50% hypodense or isodense components consistent with chronicity, as per current diagnostic standards (13).

Surgical intervention was indicated for symptomatic patients with cSDH thickness ≥10 mm on axial CT imaging. While midline shift was not an absolute criterion, its presence (particularly >5 mm) provided additional support for surgical decision-making. The final surgical decision was made by the attending neurosurgeon based on the combination of clinical symptoms, imaging findings, and the patient's overall condition (13).

Patients were included if they had a confirmed diagnosis of cSDH, based on the clinical and radiological criteria described above, and if preoperative and follow-up CT scans were available. Patients were excluded if they had undergone previous cranial surgery or had received middle meningeal artery embolization. Since death data were available from government registries, patients with incomplete clinical follow-up were still included in survival analyses.

Recurrence was defined as the clinical reappearance of neurological symptoms in a previously treated patient requiring repeat surgical intervention, as determined by the treating neurosurgeon. Asymptomatic radiological reaccumulation without clinical symptoms was not classified as recurrence, consistent with the clinical relevance-based approach to cSDH management (13).

Clinical Variables

For each patient, we recorded the following clinical variables: age, sex, American Society of Anesthesiologists (ASA) score, preoperative modified Rankin Scale (mRS), side of the operated hematoma, surgical technique (categorized as burr hole or keyhole), length of hospital stay, living status (1: home-care for self, 2: home-with assistance, 3: institutionalized), and various laboratory parameters (including hemoglobin, leukocyte count, and albumin levels). Survival data, defined as the time from the date of surgery to death from any cause, were obtained from our institutional records and verified through national registries.

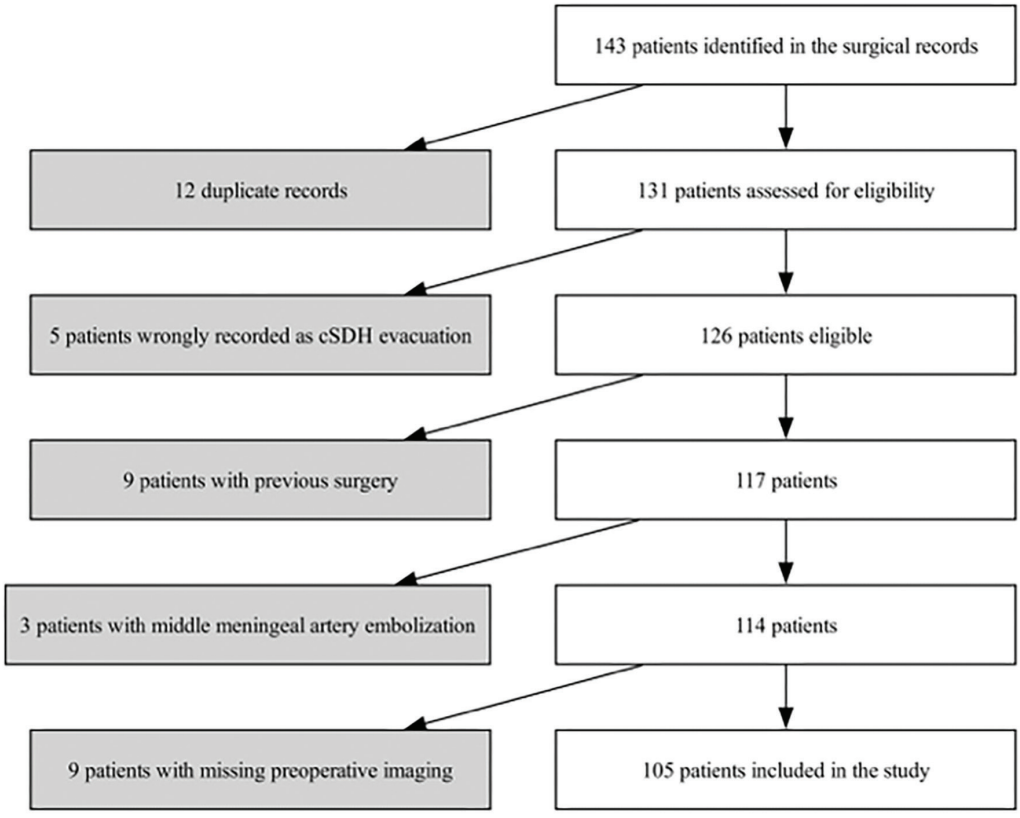


Figure 1. Patient selection flowchart
cSDH: Chronic subdural hematoma

Surgical Technique

At our center, we employ a standardized surgical approach for cSDH drainage, with technique-specific variations. All procedures were performed under general anesthesia in the supine position. For double-burr hole procedures, two separate skin incisions were made, and two burr holes (14 mm diameter each) were created. For keyhole procedures, a single skin incision (approximately 3x3 cm) was made to perform a keyhole craniotomy. Each operated side was irrigated with a minimum of 1 liter of body-temperature 0.9% NaCl solution or until the effluent was clear. Irrigation pressure was standardized across both techniques. This practice, applied uniformly throughout the study period, is concordant with subsequent randomized evidence and current guideline recommendations favoring warmed irrigation fluid (12,14). The size of the craniotomy is shown in Figure 2.

In double burr-hole procedures, flexible Penrose drains were placed and secured to the scalp, whereas keyhole surgeries used a closed subdural drainage system (Jackson-Pratt drain). Both techniques included standard closure, including galea suturing and skin approximation, and both drains were typically removed on postoperative days 2-3.

Selection of surgical technique was determined by the preferences of individual attending surgeons, with each surgeon using a consistent technique across their cases regardless of patient age, comorbidity burden, hematoma characteristics, or neurological status.

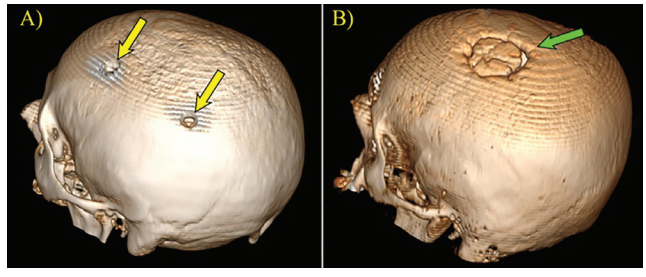


Figure 2. Three-dimensional CT reconstructions demonstrating postoperative skull appearances following the two surgical techniques used for chronic subdural hematoma drainage, viewed from the posterolateral perspective of the left side. (A) Burr hole craniotomy showing two small circular defects (yellow arrows). (B) Keyhole craniotomy demonstrating a larger oval craniotomy defect (green arrow), which allows wider surgical access and direct visualization of the subdural space
CT: Computed tomography

Temporal Muscle Thickness Measurement

Temporal muscle thickness was measured on preoperative CT scans and repeated on one-month postoperative control scans using the same technique described by Korhonen et al. (8) and Pesonen et al. (11). Two neurosurgeons, blinded to patient outcomes, independently measured TMT on both the left and right sides; the mean of these bilateral measurements was then calculated and used for subsequent analysis.

Temporal muscle thickness categories were defined using a regression-based approach. Preoperative TMT was modelled by ordinary least-squares linear regression in our cohort, first with sex alone (sex-adjusted model) and then with age and sex (age- and sex-adjusted model). Patients whose measured TMT fell below the value predicted by the respective model were classified as "Thin" and those at or above it as "Thick" The sex-adjusted classified as "Thin," and those at or above it as "Thick." The sex-adjusted classification was used for the first Kaplan-Meier comparison, and the age- and sex-adjusted classification for the second Kaplan-Meier comparison, the cohort characteristics, and the recurrence analysis. Separately, following the methodology of Korhonen et al. (8), sex-specific TMT cut-off values for predicting two-year mortality were determined using receiver operating characteristic curves by maximizing Youden's index. These cut-offs are reported as exploratory and were not used to define Thin/Thick groups in any survival or recurrence analysis.

Statistical Analysis

Statistical analyses were performed using Python 3 (Spyder IDE 6.0.3) and specialized libraries, including Pandas, NumPy, SciPy, statsmodels, and lifelines. The relevant code and dataset can be found at "https://github.com/AkbulutBB/Statistics/tree/main/cSDHTMT."

Descriptive statistics were computed for all continuous and categorical variables. Group comparisons were conducted using Student's t-tests for continuous variables and chi-square or Fisher's exact tests for categorical data. Kaplan-Meier survival analysis was employed to estimate survival probabilities at 3, 12, 24, and 60 months, and log-rank tests were used to compare survival curves between groups. Forest plots were generated to visually represent HR and confidence intervals from Cox proportional hazards models.

Multivariate Cox regression models were constructed to identify independent predictors of overall mortality and hematoma recurrence and to evaluate the prognostic impact of preoperative TMT (and its percent change at 1 month) and hematoma laterality, which was modeled categorically (right-sided hematomas as the reference). Candidate covariates were selected a priori on the basis of established clinical relevance and prior literature (age, sex, ASA score, albumin, living status, hematoma laterality, and surgical technique) and were supplemented by variables associated with mortality on univariate analysis ($p<0.10$). This combined clinical-statistical selection strategy was adopted to balance model parsimony against the limited number of observed events. Additionally, linear regression analyses were performed to assess predictors of TMT (adjusting for age, sex, ASA score, operated hematoma side, and preoperative mRS) and to examine factors affecting functional recovery by modeling mRS improvement; baseline mRS was included to address ceiling effects.

To evaluate potential selection bias in surgical technique assignment, a logistic regression analysis was performed to assess whether patient frailty markers (age, ASA score, preoperative mRS, and living status) influenced the choice between keyhole and burr hole approaches. This analysis was necessary to confirm that surgical technique selection was based solely on surgeon preference rather than on patient characteristics, which could otherwise confound the observed survival differences.

Interrater reliability for TMT measurements was quantified using intraclass correlation coefficients (ICCs) based on a two-way mixed-effects model.

A $p<0.05$ was considered statistically significant throughout the analyses.

Results

A total of 105 patients were included. The median observation time was 55.0 months (interquartile range: 30.0-91.0), and 46.7% were followed for 5 years or longer. Cohort characteristics are presented in Table 1. Interrater reliability for TMT measurements was excellent (ICC >0.9). Overall survival was 96.2% at 3 months, 91.4% at 12 months, 86.6% at 24 months, and 63.7% at 60 months; 41 deaths (39.0%) were observed.

To evaluate the influence of demographic factors on TMT, we constructed linear regression models. A sex-adjusted model explained 6.9% of the variance in TMT (R^2=0.069, p=0.007), with males exhibiting significantly greater TMT than females (β=+0.99 mm, 95% CI: 0.28 -1.71, p=0.007). Incorporating age improved the model fit (R^2=0.165, p<0.001), with older age associated with reduced TMT (β=-0.035 mm/year, p=0.001) and male sex remaining a significant predictor (β=+0.918 mm, p=0.009) (Figure 3).

Kaplan-Meier survival analysis of sex-adjusted TMT categories ("Thin" vs. "Thick") showed no significant difference in survival (log-rank p=0.149); 60-month survival rates were 71.9% for the "Thick" group and 57.2% for the "Thin" group (Figure 4). However, after further adjustment for age, the survival disparity between the groups approached statistical significance (log-rank p=0.053).

Table 1. Patient characteristics

Characteristic		Two-year mortality		Overall mortality		Total (n=105)	Unadjusted HR (95% CI)
		Alive (n=91)	Deceased (n=14)	Alive (n=64)	Deceased (n=41)		
Age	Mean (SD)	68.1 (14.4)	79.6 (10.2)	65.2 (14.6)	76.5 (11.3)	69.6 (14.4)	1.07 (1.03-1.10)*
ASA score	Mean (SD)	2.3 (0.9)	2.7 (1.0)	2.3 (0.9)	2.6 (0.8)	2.4 (0.9)	1.43 (0.99-2.07)
TMT (mm)	Mean (SD)	5.8 (1.5)	4.4 (1.9)	6.0 (1.6)	5.0 (1.6)	5.6 (1.6)	0.69 (0.54-0.89)*
Midline shift (mm)	Mean (SD)	6.6 (4.5)	7.2 (5.3)	6.4 (4.6)	7.3 (4.6)	6.7 (4.6)	1.04 (0.97-1.11)
cSDH thickness (mm)	Mean (SD)	20.3 (7.6)	22.9 (11.5)	19.8 (7.8)	22.0 (8.7)	20.6 (8.2)	1.04 (1.00-1.08)
Hemoglobin (g/dL)	Mean (SD)	12.8 (1.9)	13.1 (8.1)	12.9 (1.9)	12.7 (4.9)	12.8 (3.4)	0.95 (0.82-1.11)
Albumin (g/L)	Mean (SD)	41.4 (4.0)	36.6 (6.0)	41.5 (4.2)	39.5 (5.0)	40.7 (4.6)	0.89 (0.83-0.96)*
Male sex	n (%)	71 (78.0)	8 (57.1)	52 (81.2)	27 (65.9)	79 (75.2)	0.54 (0.28-1.03)
Thin TMT (age- and sex-adjusted)	n (%)	43 (47.3)	11 (78.6)	28 (43.8)	26 (63.4)	54 (51.4)	1.86 (0.98-3.51)
Bilateral hematoma	n (%)	19 (20.9)	5 (35.7)	15 (23.4)	9 (22.0)	24 (22.9)	1.01 (0.48-2.11)
Burr hole surgery	n (%)	55 (60.4)	9 (64.3)	32 (50.0)	32 (78.0)	64 (61.0)	2.15 (1.02-4.52)*

Continuous variables are presented as mean (standard deviation); categorical variables are presented as n (%). Two-year mortality denotes death within 24 months of surgery. Unadjusted HR were obtained from univariable Cox proportional hazards models for overall mortality, each fitted with a single covariate; continuous variables were modelled per one-unit increase in the covariate. Between-group differences in the descriptive columns were assessed with Student's t-test for normally distributed continuous variables, the Mann-Whitney U test for non-normally distributed variables, and the chi-square or Fisher's exact test for categorical variables. Thin TMT was defined as measured TMT values below those predicted by a linear model including age and sex; values at or above the predicted value were classified as Thick, *: p<0.05
ASA: American Society of Anesthesiologists, CI: Confidence interval, HR: Hazard ratio, TMT: Temporal muscle thickness, cSDH: Chronic Subdural Hematoma, SD: Standard deviation

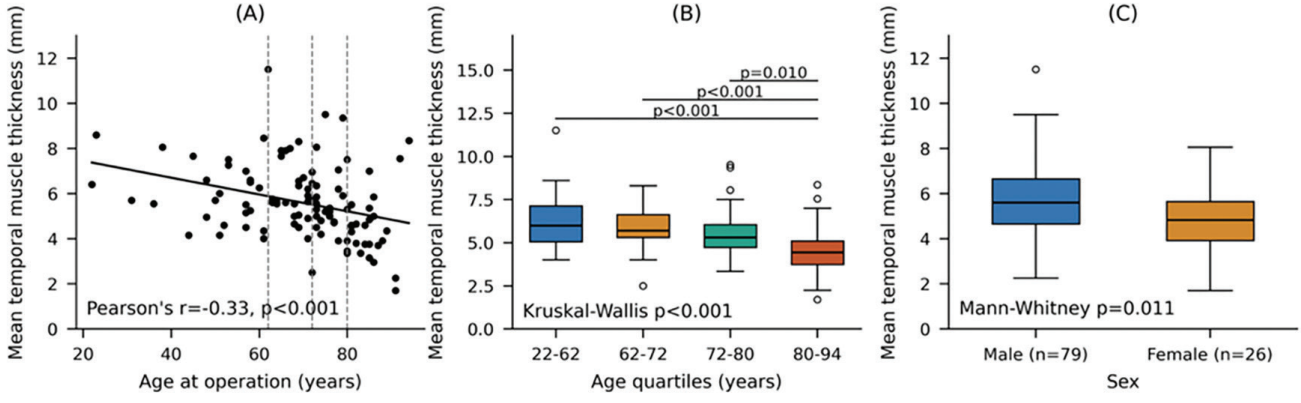


Figure 3. (A) Shows the negative correlation between age and mean temporal muscle thickness in cSDH patients. The scatterplot demonstrates a significant inverse relationship (Pearson's r=-0.33, p<0.001), indicating that TMT decreases with increasing age. (B) Displays TMT distribution across four age quartiles using boxplots. The Kruskal-Wallis test confirms significant differences between age groups (p<0.001), with pairwise comparisons showing that the oldest quartile differs significantly from each of the three younger quartiles (p<0.001, p<0.001, and p=0.010). (C) Compares TMT between male and female patients, illustrating that male patients have significantly thicker temporal muscles than female patients (Mann-Whitney p=0.011)
cSDH: Chronic Subdural Hematoma, TMT: Temporal Muscle Thickness

The age- and sex-adjusted "Thick" group demonstrated higher survival probabilities at all time points: 98.0% vs. 94.4% at 3 months; 96.1% vs. 87.0% at 12 months; 94.1% vs. 79.6% at 24 months; and 74.6% vs. 54.1% at 60 months. Median survival time in the "Thick" group was not reached, compared with 83.0 months in the "Thin" group (Figure 5).

To examine the relationship between frailty markers and TMT, a multivariable linear regression model was constructed, including age, sex, ASA score, preoperative mRS, living status, albumin levels, leukocyte counts, and hemoglobin levels (Table 2). The model accounted for 30.1% of the variance in TMT (R^2=0.301, p<0.001). Frailty markers were significantly correlated with TMT, with older age, female sex, and dependent living status predicting reduced TMT, whereas higher ASA scores were associated with greater TMT. Notably, albumin levels and hematological parameters did not reach statistical significance in this

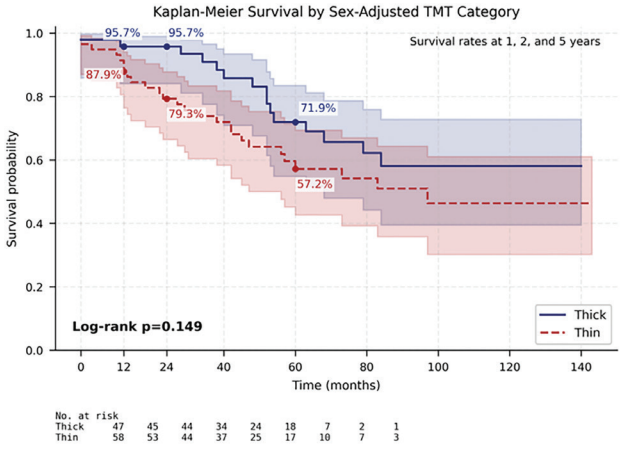


Figure 4. Kaplan-Meier survival curves for patients with chronic subdural hematoma, stratified by sex-adjusted temporal muscle thickness categories. Patients were classified as having "Thick" (solid blue line) or "Thin" (dashed red line) temporal muscles based on a sex-adjusted linear regression model. The 60-month survival rate was 71.9% in the "Thick" group and 57.2% in the "Thin" group. The shaded areas represent 95% confidence intervals. Despite this difference in survival rates, the curves did not reach statistical significance (log-rank p=0.149)
TMT: Temporal muscle thickness

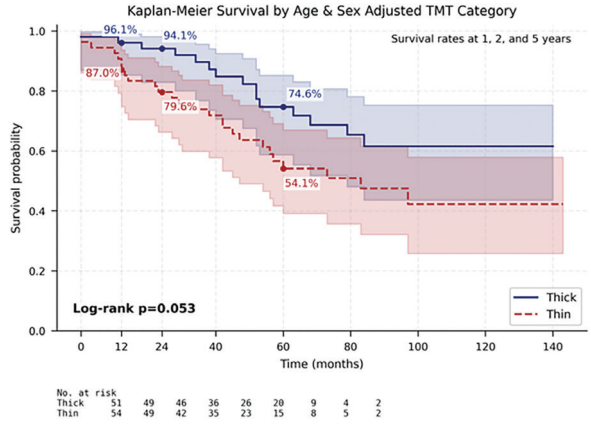


Figure 5. Kaplan-Meier survival curves for patients with chronic subdural hematoma, stratified by age- and sex-adjusted temporal muscle thickness categories. The solid blue line represents patients with "Thick" temporal muscles, while the dashed red line represents patients with "Thin" temporal muscles, after adjustment for both age and sex. Shaded areas represent 95% confidence intervals. The difference between the groups approached statistical significance (log-rank p=0.053). Survival rates at key time points (1, 2, and 5 years) are shown on the graph for both groups
TMT: Temporal muscle thickness

Table 2. Multivariable linear regression model for temporal muscle thickness

Variable	Coefficient (β)	Standard error	95% CI	p-value
Demographic factors				
Age (years)	-0.024	0.011	-0.046 to -0.002	0.032*
Sex (male vs. female)	0.765	0.340	0.090 to 1.441	0.027*
Frailty markers				
ASA score	0.554	0.203	0.150 to 0.957	0.008*
Preoperative mRS	-0.127	0.150	-0.424 to 0.170	0.399
Living status	-0.936	0.366	-1.663 to -0.210	0.012*
Nutritional markers				
Albumin (g/L)	0.070	0.036	-0.002 to 0.141	0.055
Hematological parameters				
Leukocyte count (×10³/μL)	0.001	0.054	-0.107 to 0.108	0.988
Hemoglobin (g/dL)	0.042	0.045	-0.047 to 0.130	0.351

Ordinary least-squares multivariable linear regression with TMT (mm) as the dependent variable [n=105; R^2=0.301, adjusted R^2=0.243; F(8,96)=5.18, p<0.001]. Coefficients express the change in TMT in millimetres per one-unit increase in the predictor; sex is coded male=1, female=0, and living status is coded as an ordinal score of increasing dependency, *: p<0.05
CI: Confidence interval, ASA: American Society of Anesthesiologists, mRS: modified Rankin scale, TMT: Temporal muscle thickness

comprehensive model. Receiver operating characteristic curve analysis identified sex-specific TMT cut-off values for predicting two-year mortality: 5.625 mm for males [area under the curve (AUC) 0.68; sensitivity 87.5%, specificity 52.1%] and 3.450 mm for females (AUC 0.79; sensitivity 66.7%, specificity 100%) (Table 3).

Two multivariate Cox proportional hazards models were constructed to assess predictors of overall mortality (Table 4). Model A included TMT alongside traditional clinical variables, while Model B excluded TMT to evaluate model stability. Both models demonstrated excellent concordance (0.81) and identified consistent independent predictors of mortality. Age, lower albumin levels, dependent living status, keyhole surgical technique, and bilateral hematoma location emerged as significant prognostic factors. Notably, TMT did not reach statistical significance as an independent predictor after adjustment for established clinical markers.

Table 3. Sex-specific temporal muscle thickness cut-off values for two-year mortality						
Sex	**n**	**Two-year deaths**	**Cut-off (mm)**	**Youden J**	**AUC**	**Sensitivity/specificity (%)**
Male	79	8	5.625	0.396	0.680	87.5/52.1
Female	26	6	3.450	0.667	0.792	66.7/100.0

Cut-off values were derived from receiver operating characteristic curve analysis of preoperative TMT for prediction of death within 24 months by maximizing Youden's index; calculations were performed separately by sex. Candidate thresholds were set at midpoints between consecutive observed values. Given the small number of two-year deaths within each sex stratum, these thresholds are exploratory and have not been externally validated
AUC: Area under the curve, TMT: Temporal muscle thickness

Table 4. Multivariable Cox regression models for overall mortality						
Variable	**Model A (with TMT)**			**Model B (without TMT)**		
	HR	**95% CI**	**p-value**	**HR**	**95% CI**	**p-value**
Age (per year)	1.06	1.02-1.10	<0.005*	1.07	1.03-1.11	<0.005*
Male sex	0.53	0.24-1.19	0.12	0.50	0.23-1.09	0.08
ASA score	0.81	0.45-1.44	0.47	0.78	0.44-1.39	0.41
Preoperative mRS	1.26	0.88-1.80	0.21	1.28	0.90-1.82	0.18
Hematoma thickness (mm)	1.02	0.97-1.08	0.43	1.02	0.97-1.08	0.41
Midline shift (mm)	1.04	0.95-1.14	0.36	1.04	0.95-1.14	0.39
Albumin (g/L)	0.92	0.84-1.00	0.05	0.91	0.84-0.99	0.03*
Leukocyte count ($\times 10^3/\mu L$)	1.10	0.96-1.25	0.17	1.10	0.96-1.25	0.17
Hemoglobin (g/dL)	0.97	0.80-1.17	0.74	0.96	0.79-1.17	0.71
Keyhole vs. burr hole	0.18	0.07-0.47	<0.005*	0.18	0.07-0.47	<0.005*
Living status	2.58	1.07-6.24	0.04*	2.66	1.10-6.42	0.03*
Left vs. right	0.75	0.32-1.76	0.50	0.73	0.31-1.71	0.47
Bilateral vs. right	3.23	1.22-8.55	0.02*	3.17	1.20-8.33	0.02*
TMT (mm)	0.94	0.72-1.21	0.61	-	-	-

Multivariable Cox proportional hazards regression (n=105; 41 deaths). Model A includes TMT as a covariate (concordance 0.81; likelihood-ratio test p<0.001); Model B excludes TMT (concordance 0.81; likelihood-ratio test p<0.001). Both models were adjusted for age, sex, ASA score, preoperative mRS, hematoma thickness, midline shift, albumin, leukocyte count, hemoglobin, surgical technique, living status, and hematoma laterality. Right-sided hematoma was the reference category for laterality, and burr hole craniostomy was the reference category for surgical technique. Continuous variables are modelled per one-unit increase, *: p<0.05
ASA: American Society of Anesthesiologists, CI: Confidence interval, HR: Hazard ratio, mRS: Modified Rankin scale, TMT: Temporal muscle thickness

A separate Cox regression model evaluating predictors of hematoma recurrence (n=17 events, overall rate of 16.2%) identified age (HR=1.06, 95% CI: 1.00-1.11, p=0.04) and midline shift (HR=1.16, 95% CI: 1.00-1.33, p=0.04) as significant predictors. Subgroup recurrence rates were as follows: right-sided, 13.9% (5/36); left-sided, 22.2% (10/45); and bilateral 8.3% (2/24). The age- and sex-adjusted TMT category was not associated with recurrence [Thick: 19.6% (10/51) vs. Thin: 13.0% (7/54), p=0.431]. Recurrence rates were comparable between the surgical techniques (burr hole: 20.3% vs. keyhole: 9.8%, p=0.246).

To confirm that surgical technique selection was not systematically influenced by patient characteristics, logistic regression analysis was performed using frailty markers (age, ASA score, preoperative mRS, and living status) as predictors of technique choice. The model demonstrated a poor fit (Pseudo R^2=0.032, p=0.343), indicating that technique selection reflected surgeons' preferences rather than patient-specific factors.

A Cox model incorporating preoperative TMT and the percent change in TMT at one month was examined. In this model, age (HR=1.05, 95% CI: 1.01-1.09, p=0.019), keyhole surgical technique (HR=0.22, 95% CI: 0.08-0.59, p=0.003), and dependent living status (HR=2.98, 95% CI: 1.18-7.56, p=0.021) remained significant predictors of mortality, whereas neither preoperative TMT (p=0.71) nor its percent change (p=0.10) contributed significantly.

In a linear regression model predicting mRS improvement among 90 patients (15 dropped out due to missing data), the overall model was significant [R^2=0.465, F(6,83)=12.01, p<0.001]. Baseline mRS was strongly associated with improvement (β=0.777, p<0.001), indicating that patients with worse baseline function had more room to improve, which effectively addresses the ceiling effect. Additionally, mRS at discharge was a

significant predictor (β=-0.524, p<0.001), with higher immediate postoperative mRS scores associated with less subsequent improvement. In contrast, age, sex, ASA score, and TMT were not significantly associated with mRS improvement (all p>0.28).

Discussion

This single-center retrospective cohort study evaluated TMT as a prognostic marker in surgically treated cSDH over a 10-year period. While TMT demonstrated expected demographic associations with age and sex, consistent with recent normative data (11), it failed to independently predict survival or recurrence after adjustment for established clinical variables. These findings align with emerging evidence questioning the independent clinical utility of TMT in neurological conditions (9,10) and contribute to the ongoing debate regarding its role in cSDH prognostication.

The observed age- and sex-related variation in TMT is consistent with the normative reference values established by Pesonen et al. (11), who demonstrated a progressive decline in TMT from the fourth decade onward, with males consistently maintaining higher values than females across all age groups. Our finding that these demographic factors accounted for 16.5% of TMT variance underscores a critical methodological point: studies that do not adjust for age and sex when interpreting TMT measurements risk attributing age-related muscle loss to disease-specific processes, potentially inflating TMT's apparent prognostic utility.

The failure of TMT to independently predict survival after multivariate adjustment is consistent with findings from Yilmaz et al. (9) and Atalay et al. (10), who similarly demonstrated that TMT associations with clinical outcomes were largely attenuated after adjustment for age and comorbidities across different neurological populations. The marginal survival trend observed in age- and sex-adjusted TMT categories (log-rank p=0.053) is intriguing but falls short of statistical significance and likely reflects the underpowered nature of this comparison rather than indicating a true null effect, as confirmed by our post-hoc power analysis.

The prognostic dominance of age, albumin, living status, and hematoma laterality over TMT reflects the multifactorial nature of cSDH outcomes. The independent association of albumin with mortality aligns with established evidence linking nutritional status to surgical outcomes in elderly neurosurgical patients (15,16) and supports the broader argument that systemic physiological reserve, rather than any single imaging-derived morphometric parameter, is the primary determinant of survival in this population. The significant mortality risk associated with bilateral hematomas likely reflects greater intracranial burden and a more severe underlying coagulopathy or atrophy-related vulnerability. The absence of a hemoglobin effect is consistent with external validation data from Holl et al. (17), suggesting that hematological parameters add limited independent prognostic value once nutritional and functional status are accounted for.

An unexpected association was observed between the keyhole surgical technique and reduced mortality (HR=0.18, p<0.005). This finding must be interpreted with considerable caution for several reasons. First, surgical technique selection was based on individual surgeon preference rather than randomization. Although our logistic regression confirmed that frailty markers did not predict technique choice, unmeasured confounders-such as surgeon experience, intraoperative complexity, and subtle differences in patient selection criteria-cannot be excluded in a retrospective design. Second, the magnitude of this effect (HR=0.18) substantially exceeds what has been reported in the existing literature; a meta-analysis by Huang et al. (18), including 3,559 surgeries, found comparable mortality outcomes between burr hole and craniotomy techniques, and a Turkish tertiary center similarly reported no survival advantage for craniotomy (19). Current evidence has made this discordance more difficult to explain. The recent multidisciplinary guideline recommends burr-hole drainage with drain placement as first-line surgical management and reserves craniotomy for selected cases. It attributes no survival advantage to either approach (12). The multicenter COMPACT randomized trial compared mini-craniotomy, twist-drill craniostomy, and burr hole craniostomy and found no significant difference in outcome between arms; burr hole craniostomy had the lowest reoperation rate (14). Our estimate, therefore, conflicts with both observational and randomized evidence, which argue against a causal effect.

Third, our study was underpowered for subgroup analyses, and extreme HR in small retrospective cohorts are well-recognized statistical phenomena when the number of events is limited. The keyhole technique did not demonstrate a recurrence advantage over burr hole drainage (9.8% vs. 20.3%, p=0.246), further undermining a straightforward causal interpretation. Regarding drainage systems, the closed Jackson-Pratt drain used in keyhole procedures versus the open Penrose drain in burr hole procedures may theoretically influence outcomes, though studies have not demonstrated consistent infection-related advantages of closed systems (20,21). We therefore present this finding as hypothesis-generating only and strongly advocate conducting prospective randomized trials before any practice-changing conclusions are drawn.

Our recurrence rate of 16.2% falls within the 10-20% range reported after burr hole drainage (22). Our cohort was therefore not unusually low-risk, which

strengthens the finding that TMT was unrelated to recurrence. This is clinically important because recurrence imposes a substantial burden. After reoperation for a first recurrence, a further 20% of patients recur, and most are left with residual symptoms and moderate-to-poor functional recovery (22). If sarcopenia contributed to the inflammatory and angiogenic cycle that drives hematoma reaccumulation, TMT should have predicted recurrence. Recurrence was predicted by age and midline shift, both of which reflect intracranial processes rather than muscle mass. TMT, therefore, appears to index general physiological reserve rather than any hematoma-specific mechanism, and preoperative morphometry is unlikely to identify patients at risk of recurrence.

Du et al. (6) raised several important methodological concerns regarding TMT studies in cSDH, advocating for comprehensive frailty assessment beyond single morphometric parameters. Our study partially addresses these concerns by incorporating albumin as a nutritional marker, standardizing surgical protocols, and examining hematoma laterality as an independent variable. However, we acknowledge that relying on TMT and albumin alone provides an incomplete picture of patient frailty. Validated multidimensional frailty instruments, such as the Charlson Comorbidity Index or the modified frailty index, and inflammatory markers such as C-reactive protein and lymphocyte count were not available in our retrospective dataset. This gap is now more consequential because current guidance recommends screening every admitted cSDH patient with a validated frailty tool such as the clinical frailty scale and referring those who screen positive for comprehensive geriatric assessment (12). Any future claim that TMT adds prognostic information will therefore have to be tested against a frailty score already collected in routine care, rather than against age and comorbidity alone. This limitation is not unique to our study; Mohajerani et al. (23) similarly noted that multiple confounders, including age, sex, comorbidities, nutritional status, and preexisting frailty, may influence the independent prognostic value of TMT in neurological populations. Prospective studies should incorporate standardized frailty assessment tools alongside TMT to determine whether a composite frailty score provides superior prognostic discrimination compared with TMT alone.

Study Limitations

This study has several limitations that should be considered when interpreting its findings. The single-center retrospective design limits generalizability to other institutional settings, and the assignment of surgical technique was based on surgeon preference rather than randomization, introducing potential residual confounding despite our logistic regression validation. The sample size of 105 patients, while comparable to prior TMT studies in cSDH (7,8), may limit the statistical power of subgroup analyses. Post-hoc power analysis of the Thin versus Thick TMT survival comparison revealed that 41 observed events provided approximately 72% power to detect the HR of 0.48, falling short of the 57 events required for 80% power. Accordingly, the non-significant trend observed (log-rank p=0.053) may reflect insufficient statistical power rather than a true absence of effect, and the potential prognostic role of TMT cannot be definitively excluded. The relatively small number of recurrence events (n=17) further limits the recurrence analysis, and the results should be interpreted cautiously. The sex-specific TMT cut-off values derived from our cohort should be interpreted cautiously due to the relatively small sample size and the lack of external validation. Additionally, our retrospective design precluded the incorporation of validated multidimensional frailty instruments, such as the Charlson Comorbidity Index and the modified frailty index, and inflammatory markers, such as C-reactive protein and lymphocyte count, limiting the comprehensiveness of frailty assessment. The one-month postoperative TMT follow-up may also be insufficient to capture longer-term changes in muscle mass. Finally, our study period predates the adoption of middle meningeal artery embolization at our center, and the few patients who received it were excluded by design. Since randomized and pooled data now indicate that adjunctive embolization approximately halves the risk of recurrence without a demonstrable effect on mortality or functional outcome (24), our results should be read as applying to conventionally operated patients; whether TMT retains prognostic value in cohorts managed with embolization remains untested.

Despite these limitations, this study offers notable strengths. The 10-year follow-up, with a median of 55.0 months and nearly half of patients exceeding 5 years, substantially exceeds durations reported in prior TMT studies. The rigorous age- and sex-adjusted TMT methodology, the excellent interrater reliability (ICC>0.9), and the inclusion of surgical technique and hematoma laterality as previously unexamined covariates represent meaningful methodological contributions to the field.

Conclusion

Temporal muscle thickness does not independently predict survival or recurrence in surgically treated cSDH patients when traditional clinical markers are considered. Age, albumin levels, living situation, and hematoma laterality remain robust prognostic determinants. The marginal survival trend observed in age- and sex-adjusted TMT categories warrants further investigation in larger, prospective, multicenter cohorts with adequate statistical

power. The unexpected association between the keyhole technique and survival, while hypothesis-generating, requires validation through randomized controlled trials before clinical implications can be drawn.

Ethics

Ethics Committee Approval: Ethical approval was obtained from the Ege University Ethics Committee (approval number: 24-3T/34, date: 07.03.2024).

Informed Consent: The patients consent to participate was obtained using the ethics committee-approvaed form.

Footnotes

Authorship Contributions

Concept: H.B., E.O., T.Y., Design: B.B.A., E.E.C., Data Collection or Processing: B.B.A., E.E.C., M.S.B., Analysis or Interpretation: B.B.A., M.S.B., Literature Search: B.B.A., E.E.C., Writing: B.B.A., E.E.C., M.S.B., H.B., E.O., T.Y.

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

Declaration Regarding the Use of AI and AI-Assisted Technologies: During the preparation of this work, the authors used Claude Sonnet 4 (Anthropic, San Francisco, CA, USA) to check the grammar and improve the flow of sentences. After using this tool, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

References

1. Dziho A, Al Awadhi A, Mollica C, et al. Global prevalence and incidence of chronic subdural hematoma: a systematic review. Brain Spine. 2025;5:105893.
2. Zhu F, Wang H, Li W, et al. Factors correlated with the postoperative recurrence of chronic subdural hematoma: an umbrella study of systematic reviews and meta-analyses. EClinicalMedicine. 2021;43:101234.
3. Nouri A, Gondar R, Schaller K, Meling T. Chronic subdural hematoma (cSDH): a review of the current state of the art. Brain Spine. 2021;1:100300.
4. Posti JP, Luoto TM, Sipilä JOT, Rautava P, Kytö V. Prognosis of patients with operated chronic subdural hematoma. Sci Rep. 2022;12:7020.
5. Furtner J, Weller M, Weber M, et al. Temporal muscle thickness as a prognostic marker in patients with newly diagnosed glioblastoma: translational imaging analysis of the CENTRIC EORTC 26071–22072 and CORE trials. Clin Cancer Res. 2022;28:129-36.
6. Du X, Chen G, Ren Z. Comment on: 'reduced temporal muscle thickness predicts shorter survival in patients undergoing chronic subdural haematoma drainage' by Korhonen et al. J Cachexia Sarcopenia Muscle. 2024;15:2879-80.
7. Dubinski D, Won SY, Behmanesh B, et al. Significance of temporal muscle thickness in chronic subdural hematoma. J Clin Med. 2022;11:6456.
8. Korhonen TK, Arponen O, Steinruecke M, et al. Reduced temporal muscle thickness predicts shorter survival in patients undergoing chronic subdural haematoma drainage. J Cachexia Sarcopenia Muscle. 2024;15:1441-50.
9. Yilmaz E, Arsava EM, Topcuoglu MA. Temporalis muscle thickness: ultrasound measurement, clinical significance and correlation with sarcopenic indices. Clin Nutr. 2025;54:140-52.
10. Atalay B, Erincik K, Dogan MB, Gezgin M, Yildiz H, Ozdilek B. The impact of temporal muscle thickness as an indicator of sarcopenia on clinical status in Parkinson's disease. Clin Neurol Neurosurg. 2026;265:109388.
11. Pesonen EK, Arponen O, Niinimäki J, et al. Age- and sex-adjusted CT-based reference values for temporal muscle thickness, cross-sectional area and radiodensity. Sci Rep. 2025;15:2393.
12. Stubbs DJ, Davies BM, Edlmann E, et al. Clinical practice guidelines for the care of patients with a chronic subdural haematoma: multidisciplinary recommendations from presentation to recovery. Br J Neurosurg. 2026;40:94-103.
13. Kan P, Fiorella D, Dabus G, et al. ARISE i consensus statement on the management of chronic subdural hematoma. Stroke. 2024;55:1438-48.
14. Fakhry R, Yesildal C, Bartek J, et al. Updated systematic review of current randomised controlled trials in chronic subdural haematoma. Acta Neurochir (Wien). 2025;167:288.
15. Wang S, Ma Y, Zhao X, et al. Risk factors of hospital mortality in chronic subdural hematoma: a retrospective analysis of 1117 patients, a single institute experience. J Clin Neurosci. 2019;67:46-51.
16. Zhu B, Yu X, Ou Y, Guo X, Liu W, Wu L. Nutritional and inflammatory peripheral blood markers for risk assessment of chronic subdural hematoma: a case-control study. Clin Neurol Neurosurg. 2023;227:107640.
17. Holl DC, Mikolic A, Blaauw J, et al. External validation of prognostic models predicting outcome after chronic subdural hematoma. Acta Neurochir (Wien). 2022;164:2719-30.
18. Huang YW, Yin XS, Li ZP. Burr hole craniostomy vs. minicraniotomy of chronic subdural hematoma: a systematic review and meta-analysis. Eur Rev Med Pharmacol Sci. 2022;26:4983-90.
19. Hamzaoglu V, Özalp H, Avcı E, Bağdatoğlu C, Dağtekin A. Surgical treatment outcomes of our patients with chronic subdural hematoma: a review of seventy six cases. Ankara Üniversitesi Tıp Fakültesi Mecmuası. 2018;71:251-55.
20. Bakhti S, Tighilt N, Djennas M. Open drainage in chronic subdural hematomas: a prospective study of 189 cases. Iran J Neurosurg. 2016;2:15-9.
21. Morales-Gómez JA, Garcia-Estrada E, Garza-Báez A, Mercado-Flores M, de León AMP. Subdural open drains as an effective and low-cost modality for the treatment of chronic subdural hematomas. Br J Neurosurg. 2023;37:1078-81.
22. de Rooij MJ, van der Boog ATJ, Carter SH, Robe PAJT, van der Zwan A, Niknejad HR. Clinical outcome of redo-surgery for

recurrent chronic subdural hematoma: a systematic review and meta-analysis combined with a case series. World Neurosurg. 2025;202:124377.

23. Mohajerani E, Gümüs M, Dinger TF, et al. Temporal muscle thickness as a feasible sarcopenia marker and outcome predictor after aneurysmal subarachnoid hemorrhage. Acta Neurochir (Wien). 2025;167:57.

24. Mohammadzadeh I, Hajikarimloo B, Zare A, et al. Efficacy of middle meningeal artery embolization combined with surgery versus standalone surgery for chronic subdural hematoma: a comprehensive systematic review and meta-analysis with separate analysis of randomized controlled trials. Neuroradiology. 2025;67:2897-919.

Original Article

DOI: 10.4274/haseki.galenos.2026.57070
Med Bull Haseki 2026;64(5):426-434

Clinical Predictors of Proximal Esophageal Impaction and Their Implications for Management in Pediatric Foreign Body Ingestion: A Five-year Retrospective Cohort Study

Mehmet Necmeddin Sutasir[1], Ahmet Bas[2], Betul Beyza Ozbal[2], Fatma Sarac[2], Banu Arslan[1]

[1]*University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Emergency Medicine, Istanbul, Türkiye*
[2]*University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Pediatric Surgery, Istanbul, Türkiye*

Abstract

Aim: The anatomical level of esophageal impaction plays a pivotal role in determining the optimal management strategy in children with foreign-body ingestion. In this context, we aimed to evaluate the clinical characteristics, management strategies, and outcomes of pediatric patients presenting with foreign-body ingestion to a tertiary-care pediatric emergency department, with particular emphasis on factors associated with the level of esophageal impaction.

Methods: This retrospective, single-center cohort study included 397 consecutive pediatric patients (0-18 years) evaluated for suspected foreign-body ingestion in the Pediatric Emergency Clinic of University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital between July 2020 and June 2024. Patient demographics, foreign body characteristics, anatomical level of esophageal impaction, management strategies, extraction techniques, and clinical outcomes were analyzed. Independent predictors of proximal esophageal impaction were identified using multivariable logistic regression, and the discriminatory performance of the final model was evaluated by receiver operating characteristic curve analysis.

Results: A total of 397 pediatric patients were included [220 males (55.4%), 150 females (37.8%); mean age, 49.1±42.5 months]. Children aged 12-36 months represented the largest age group (35.3%). Coins were the most frequently ingested foreign bodies (56.4%), and the cricopharyngeal region was the predominant site of esophageal impaction (63.5%). Although coin ingestion showed a higher rate of proximal impaction than other foreign bodies in univariate analysis (80.9% vs. 63.3%), the association did not reach statistical significance (p=0.053). After adjustment for potential confounders, coin ingestion [odds ratio (OR)=2.41, 95% confidence interval (CI): 1.52-3.82; p<0.001] and age ≤36 months (OR=1.68, 95% CI: 1.08-2.61; p=0.021) emerged as independent predictors of first physiological esophageal narrowing impaction. The multivariable model demonstrated acceptable discriminative performance (area under the curve =0.71, 95% CI: 0.65-0.77; p<0.001). Overall, successful extraction was achieved in 84.4% of patients, with an overall complication rate of 0.25%.

Conclusion: Younger age and coin ingestion independently predicted proximal esophageal impaction. The proposed model demonstrated acceptable discrimination and may support clinical decision-making in pediatric emergency practice.

Keywords: Foreign bodies, child, esophagus, endoscopy, emergency service

Corresponding Author: Mehmet Necmeddin Sutasir, MD, University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Emergency Medicine, Istanbul, Türkiye
E-mail: drmehmetns@gmail.com **ORCID:** orcid.org/0000-0001-6472-5092
Received: 12.02.2026 **Accepted:** 17.07.2026 **Publication Date:** 17.11.2026

Cite this article as: Sutasir MN, Bas A, Ozbal BB, Sarac F, Arslan B. Clinical predictors of proximal esophageal impaction and their implications for management in pediatric foreign body ingestion: a five-year retrospective cohort study. Med Bull Haseki. 2026;64(5):426-434

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

Introduction

Foreign body ingestion (FBI) is one of the most common pediatric emergencies in emergency departments worldwide and remains an important cause of urgent gastrointestinal intervention. The vast majority of cases occur in children younger than five years because of developmentally appropriate oral exploration, immature swallowing coordination, and limited hazard awareness. Although most ingested foreign bodies pass spontaneously through the gastrointestinal tract, esophageal impaction may lead to serious complications, including mucosal injury, ulceration, perforation, mediastinitis, and, in selected cases, life-threatening airway compromise. Contemporary international guidelines emphasize that management should be individualized according to four key factors: the type of ingested object, its anatomical location, the patient's clinical condition, and the time elapsed since ingestion (1-3).

Coins remain the most commonly ingested foreign bodies in children, accounting for the majority of esophageal foreign-body presentations in many countries. Because coin dimensions vary across healthcare settings, local epidemiological data remain clinically important. The diameters of circulating coins in Türkiye range from 17.0 to 26.15 mm; this variation may influence the likelihood of esophageal impaction depending on the child's age and esophageal caliber. The present data also indicates that the morphology, size, and anatomic orientation of the objects are risk factors for impaction, whereas button batteries, magnets, and sharp objects are associated with different management strategies owing to their significantly higher complication rates (1,4-6).

The management of pediatric esophageal foreign bodies includes various therapeutic options, such as observation, Magill forceps extraction, Foley catheter retrieval, rigid esophagoscopy, and flexible endoscopy. Selection of the most appropriate intervention depends on the level of impaction, object characteristics, patient stability, and local expertise. Recent guideline updates recommend a structured, risk-based approach rather than the routine use of a single extraction technique and encourage individualized treatment algorithms to optimize procedural success while minimizing complications (1,3,7,8).

We hypothesized that coin ingestion and younger age are each independently associated with proximal esophageal impaction and that the anatomical level of impaction influences the choice of extraction technique in pediatric patients with FBI. Despite recent international guideline updates, evidence integrating foreign body characteristics, anatomical impaction patterns, and real-world extraction strategies in large pediatric emergency cohorts remains limited, particularly in countries with different coin dimensions in circulation. This study addresses this knowledge gap by evaluating the clinical characteristics, anatomical distribution, management strategies, and outcomes of FBI in children seen in a high-volume tertiary pediatric emergency department over five years.

Materials and Methods

Compliance with Ethical Standards

This study was conducted in accordance with the ethical principles of the Declaration of Helsinki and the Good Clinical Practice guidelines. Ethical approval was obtained from the University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital Clinical Research Ethics Committee (decision no: 301817582, date: 20.01.2026). Because of the retrospective study design and the use of anonymized patient data, the requirement for individual informed consent was waived by the Ethics Committee.

Study Design and Setting

This single-center retrospective observational study was conducted at the Pediatric Emergency Clinic of University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Istanbul, Türkiye, between July 1, 2020, and June 30, 2024.

Study Population and Data Collection

The study population comprised 397 pediatric patients, aged 0-18 years, who met predefined inclusion criteria and were evaluated for suspected foreign-body ingestion during the study period. Patient eligibility was assessed retrospectively according to predefined inclusion and exclusion criteria. Only patients with complete clinical documentation and sufficient data for analysis were included in the final study cohort. Patients were excluded for any of the following: tracheobronchial aspiration, incomplete medical records, previous gastrointestinal surgery, known gastrointestinal anomalies, neuromuscular disorders affecting swallowing, or transfer from another institution while receiving ongoing treatment.

Because this study was based on retrospectively collected institutional archival data, the total number of initially screened patients and the exact number of excluded cases in each exclusion category could not be reliably reconstructed. To avoid reporting potentially inaccurate numbers, only patients who fulfilled all predefined eligibility criteria and had complete, evaluable data were included in the final analysis. This methodological limitation has been explicitly acknowledged to ensure transparency in patient selection.

Medical records were systematically reviewed using the hospital's information management system, including

electronic patient records, emergency department logs, endoscopy reports, radiology reports, and discharge summaries. A standardized data collection form was used to extract demographic information (age, sex, age group), clinical parameters (foreign body type, level of esophageal impaction, extraction technique), and outcomes. Radiological images were reviewed via the hospital PACS system. Radiolucent foreign bodies were identified based on clinical history, physical examination findings, and endoscopic evaluation when available, particularly in cases where radiographic imaging was non-diagnostic.

Inclusion and Exclusion Criteria

Patients were included if they were aged 0-18 years, had documented FBI or a radiologically confirmed gastrointestinal foreign body, had complete medical records, and underwent at least one radiological imaging study. Exclusion criteria encompassed tracheobronchial aspiration, incomplete medical records, transfer from other institutions while receiving ongoing treatment, known gastrointestinal anomalies (esophageal stricture, achalasia, pyloric stenosis, and intestinal malrotation), prior gastrointestinal surgery, and neuromuscular disorders affecting swallowing function.

Outcome Measures

The primary outcome of the study was to identify factors associated with proximal esophageal impaction (first physiological narrowing), with particular emphasis on patient age and foreign body type.

The secondary outcomes included the anatomical distribution of impacted foreign bodies, the extraction techniques used according to the level of impaction, procedural success, complications, and overall clinical outcomes.

Statistical Analysis

Statistical analyses were performed using SPSS version 26.0 (IBM Corp., Armonk, NY, USA). Categorical variables were expressed as frequencies and percentages, while continuous variables were presented as mean ± standard deviation. The Kolmogorov-Smirnov and Shapiro-Wilk tests assessed the normality of the distribution. Categorical comparisons were performed using chi-square or Fisher's exact tests, as appropriate. Multivariate binary logistic regression analysis was performed to identify independent predictors of proximal esophageal impaction (first narrowing vs. second/third narrowing). Variables considered clinically relevant and/or showing a univariate association with $p<0.10$ were selected as candidates for multivariable logistic regression. Results were expressed as odds ratios (ORs) with 95% confidence intervals (CIs). The discriminative performance of the logistic regression model was evaluated using receiver operating characteristic (ROC) curve analysis, with the area under the curve (AUC) calculated to assess model accuracy. An AUC of 0.7-0.8 was considered acceptable, 0.8-0.9 was considered excellent, and >0.9 was considered outstanding. Model calibration was assessed using the Hosmer-Lemeshow goodness-of-fit test. A two-tailed p-value below 0.05 was considered statistically significant. Object-specific variables, such as diameter, thickness, and morphology, were not included in the regression model to maintain model stability and avoid overfitting, given subgroup variability and incomplete data. Patient flow through eligibility assessment to final study inclusion is summarized in Figure 1.

Candidate variables for multivariable logistic regression included age group, sex, and foreign body type. Variables with a univariate p-value <0.10 were considered eligible for entry into the multivariable model. Object-specific variables, including coin denomination, object diameter, thickness, and morphology, were evaluated during model development but not entered into the final model because these data were incomplete or unavailable for several foreign body categories, resulting in substantial subgroup heterogeneity and an increased risk of model overfitting.

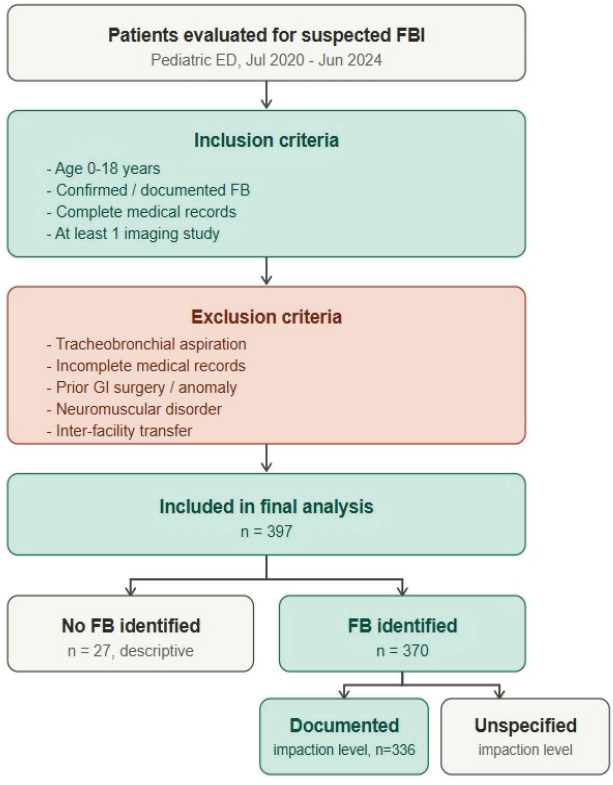


Figure 1. Study flow diagram of patient eligibility assessment and the final cohort (n=397), stratified by foreign body identification status. Screened/excluded numbers per criterion could not be reliably reconstructed from retrospective records

FBI: Foreign body ingestion, ED: Emergency department, GI: Gastrointestinal

Consequently, the final multivariable model included foreign body type (coin vs. non-coin), age group (≤36 vs. >36 months), and sex.

Results

Patient Demographics

Over the five-year study period, 397 pediatric patients were evaluated for FBI and included in the final analysis. Cases in which no foreign body was identified during the intervention (n=27) were included in descriptive analyses but excluded from comparative analyses examining relationships between foreign body type, impaction level, and extraction technique. The cohort comprised 220 males (55.4%) and 150 females (37.8%), and 27 cases (6.8%) with sex undocumented. The male-to-female ratio was 1.47:1. The mean patient age was 49.1±42.5 months (median: 36.0 months; range: 3-204 months). The 12-36-month age group accounted for the largest proportion of cases (140 patients, 35.3%), followed by the 36-60-month and 60-120-month groups (90 patients each, 22.7%) (Table 1).

Table 1. Demographic and clinical characteristics of patients (n=397)

Characteristic	n	%
Sex		
Male	220	55.4
Female	150	37.8
Unknown	27	6.8
Age group		
0-12 months	32	8.1
12-36 months	140	35.3
36-60 months	90	22.7
60-120 months	90	22.7
>120 months	35	8.8
Esophageal impaction level		
First narrowing (cricopharyngeal)	252	63.5
Second narrowing (aortic arch)	52	13.1
Third narrowing (GEJ)	32	8.1
Not specified/other	50	12.6
Extraction technique		
Rigid esophagoscopy	207	52.1
Magill forceps	108	27.2
Foley catheter	63	15.9
Flexible endoscopy	17	4.3
Procedural outcome		
Successful extraction	335	84.4
Advanced to stomach	36	9.1
No foreign body identified	27	6.8

GEJ: Gastroesophageal junction

Types of Ingested Foreign Bodies

Coins were the predominant type of foreign body, accounting for 224 cases (56.4% of the total cohort). Among coin ingestions (calculated as percentages of the coin subgroup, n=224), 1 Turkish Lira coin (diameter: 26.15 mm) was the most common (69 cases, 30.8%), followed by 25 kurus coins (50 cases, 22.3%), 50 kurus coins (39 cases, 17.4%), 10 kurus coins (14 cases, 6.3%), and 5 kurus coins (11 cases, 4.9%). The specific denomination could not be determined in 42 cases (18.8% of the coin subgroup). Button batteries represented the second most common category (33 cases, 8.3% of total), followed by organic matter (23 cases, 5.8%), needles (15 cases, 3.8%), and jewelry items (14 cases, 3.5%) (Table 2).

Esophageal Impaction Patterns

The anatomical distribution of impaction demonstrated a clear predominance at the first physiological narrowing (cricopharyngeal region), accounting for 252 cases (63.5% of the total). The second narrowing, at the level of the aortic arch, served as the impaction site in 52 cases (13.1%), while the third narrowing, at the gastroesophageal junction, was the impaction site in 32 cases (8.1%). The exact level was not specified in 50 cases (12.6%).

Relationship Between Foreign Body Type and Impaction Level

An initial univariate analysis of the relationship between foreign body type and impaction level revealed clinically noteworthy differences (χ^2=12.422, p=0.053).

Table 2. Types of ingested foreign bodies (n=397)

Foreign body type	n	%
Coins (total)	**224**	**56.4***
1 TL (26.15 mm)	69	30.8†
50 kurus (23.85 mm)	39	17.4†
25 kurus (20.5 mm)	50	22.3†
10 kurus (18.0 mm)	14	6.3†
5 kurus (17.0 mm)	11	4.9†
Unspecified denomination	42	18.8†
Button batteries	33	8.3*
Organic matter (meat, bone)	23	5.8*
Needles	15	3.8*
Jewelry (earrings, rings)	14	3.5*
Plastic/toys	5	1.3*
Magnets	4	1.0*
Other metallic objects	25	6.3*
No foreign body identified‡	27	6.8*

*Percentage calculated from total cohort (n=397). †Percentage calculated from coin subgroup (n=224) ‡Cases with clinical suspicion but no foreign body found during intervention; included in descriptive analyses only
TL: Turkish lira, mm: Millimeters

Among cases with documented impaction level (n=336, excluding those with unspecified location or no foreign body identified), coins exhibited the highest rate of impaction at the first narrowing (174/215, 80.9%), compared with those for needles (9/12, 75.0%), button batteries (19/30, 63.3%), and other objects (50/79, 63.3%). Conversely, button batteries and other objects had higher rates of impaction at the second narrowing (23.3% and 24.1%, respectively) than coins (11.2%) (Table 3).

Multivariate Logistic Regression Analysis

To further investigate predictors of proximal esophageal impaction, a multivariate binary logistic regression analysis was performed with first-narrowing impaction as the dependent variable. Variables included in the model were foreign body type (coin vs. non-coin), age group (≤36 months vs. >36 months), and sex. The 36-month cutoff was selected based on clinical relevance—children in this age group represent the highest-risk population for FBI—and on the distribution of cases within the study cohort. After adjusting for confounding variables, coin ingestion emerged as an independent predictor of impaction at the first narrowing (OR=2.41, 95% CI: 1.52-3.82, p<0.001). Age ≤36 months was also significantly associated with proximal impaction (OR=1.68, 95% CI: 1.08-2.61, p=0.021), while sex showed no significant association (OR=1.12, 95% CI: 0.72-1.74, p=0.614). Sex was not independently associated with proximal impaction (Table 4). The Hosmer-Lemeshow test indicated adequate model fit (χ^2=5.83, p=0.442) (Table 4). The ROC curve analysis demonstrated that the logistic regression model had acceptable discriminative ability, with an AUC of 0.71 (95% CI: 0.65-0.77, p<0.001). The optimal cutoff probability yielded a sensitivity of 68.3% and a specificity of 64.8% in predicting first narrowing impaction. These findings confirm that coins, with their flat, disc-like morphology that presents a larger cross-sectional area perpendicular to the esophageal lumen, are significantly more likely to lodge at the proximal esophageal narrowing than other foreign body types.

Extraction Techniques and Outcomes

Rigid esophagoscopy was the most frequently employed technique (207 cases, 52.1%), followed by Magill forceps extraction (108 cases, 27.2%), Foley catheter retrieval (63 cases, 15.9%), and flexible endoscopy (17 cases, 4.3%). The relationship between impaction level and extraction technique was statistically significant (χ^2=198.45, p<0.001). For impactions at the first narrowing, Magill forceps were most commonly employed (103/252, 40.9%), followed by rigid esophagoscopy (76/252, 30.2%) and Foley catheters (63/252, 25.0%). In contrast, impactions at the second narrowing were managed exclusively with rigid esophagoscopy (52/52, 100%), whereas impactions at the third narrowing predominantly required rigid esophagoscopy (29/32, 90.6%) or flexible endoscopy (3/32, 9.4%) (Figure 2, Table 5). Successful extraction was achieved in 335 cases (84.4%), while 36 cases (9.1%) involved intentional advancement into the stomach. One esophageal perforation (0.25%) required bilateral chest tube insertion. This patient recovered fully following an extended hospitalization.

Discussion

This study provides an integrated evaluation of pediatric FBI in a high-volume, tertiary pediatric emergency department and demonstrates that patient

Table 3. Relationship between foreign body type and esophageal impaction level (cases with specified impaction level only)

Foreign body type	First narrowing n (%)	Second narrowing n (%)	Third narrowing n (%)
Coins (n=215)	174 (80.9)	24 (11.2)	17 (7.9)
Button batteries (n=30)	19 (63.3)	7 (23.3)	4 (13.3)
Needles (n=12)	9 (75.0)	2 (16.7)	1 (8.3)
Other objects (n=79)	50 (63.3)	19 (24.1)	10 (12.7)
Total (n=336)	**252 (75.0)**	**52 (15.5)**	**32 (9.5)**

Chi-square test, χ^2=12.422, p=0.053. Cases with an unspecified location or with no foreign body identified were excluded

Table 4. Multivariate logistic regression analysis for predictors of first narrowing impaction

Variable	OR	95% CI	p-value
Coin (vs. non-coin)	2.41	1.52-3.82	**<0.001**
Age ≤36 months (vs. >36 months)	1.68	1.08-2.61	**0.021**
Male sex (vs. female)	1.12	0.72-1.74	0.614

Model fit: Hosmer-Lemeshow χ^2=5.83, p=0.442. ROC analysis: AUC=0.71 (95% CI: 0.65-0.77), sensitivity 68.3%, specificity 64.8%. Bold p-values indicate statistical significance (p<0.05)
OR: Odds ratio, CI: Confidence interval, ROC: Receiver operating characteristic, AUC: Area under the curve

age, foreign body characteristics, anatomical impaction site, and extraction strategy are closely interconnected in clinical management. Rather than merely describing the epidemiology of FBI, our findings identify clinically relevant factors associated with proximal esophageal impaction and support a structured, anatomy-based approach to selecting the most appropriate retrieval technique. These observations are consistent with recent international recommendations emphasizing that successful management depends on integrating patient characteristics with the type, location, and risk profile of the ingested object rather than relying on a single clinical parameter (1,4).

The demographic characteristics of our cohort were comparable to those reported in recent international studies. The predominance of male patients and the concentration of cases among children younger than 3 years are well recognized and are generally attributed to developmental oral exploration, immature swallowing coordination, and increased exposure to small household objects during early childhood (2,5,9-11). Recent international reviews have reinforced the epidemiology of FBI to show the influence of regional differences in household environments and the availability of commonly ingested objects, emphasizing the necessity of population-specific epidemiological data (4,11).

Coins remained the predominant ingested object in our cohort, consistent with contemporary pediatric series showing that coins continue to represent the most frequent gastrointestinal foreign body despite growing awareness of high-risk objects such as button batteries and magnets (5,6,9,11). Although current guidelines appropriately prioritize button battery ingestion because of the potential for rapid tissue injury, coin ingestion continues to generate a substantial proportion of pediatric emergency presentations (1,4,8). The predominance of the 1 Turkish Lira coin in our population probably reflects local circulation and accessibility patterns, highlighting the

Table 5. Relationship between esophageal impaction level and extraction technique

Impaction level	Rigid esophagoscopy n (%)	Magill forceps n (%)	Foley catheter n (%)	Flexible endoscopy n (%)
First narrowing (n=252)	76 (30.2)	103 (40.9)	63 (25.0)	9 (3.6)
Second narrowing (n=52)	52 (100)	0 (0)	0 (0)	0 (0)
Third narrowing (n=32)	29 (90.6)	0 (0)	0 (0)	3 (9.4)

Chi-square test, p<0.001. Cases with unspecified location (n=50) were excluded from this analysis

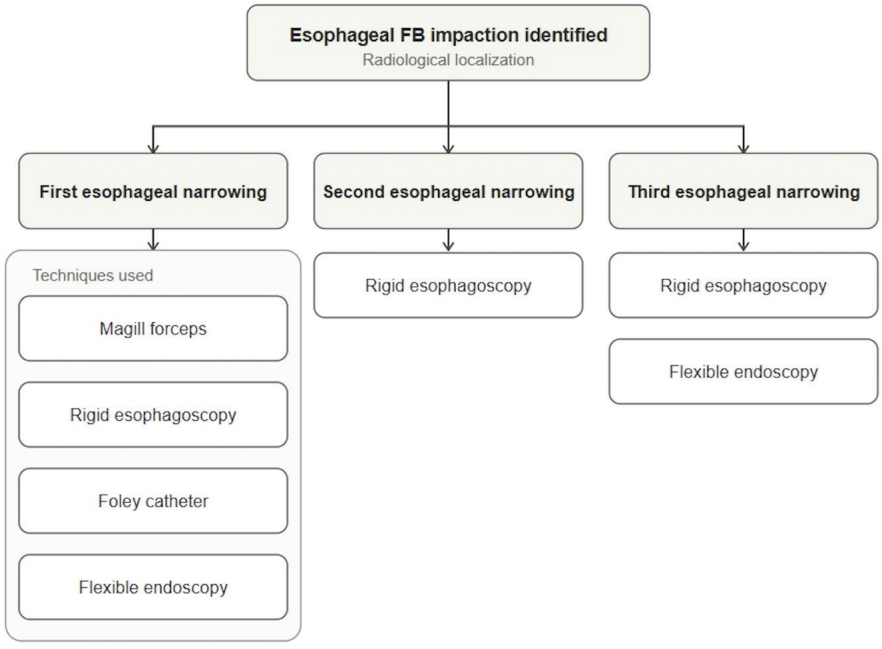


Figure 2. Management algorithm by esophageal narrowing level. Listed techniques are not sequential; selection is individualized based on object characteristics, patient stability, and institutional expertise
FB: Foreign body

importance of regional context when comparing foreign body distributions across countries (9,11).

One of the most clinically relevant observations was that coins tended to become impacted at the first physiological narrowing of the esophagus more frequently than other objects. This pattern is biologically plausible in the setting of esophageal anatomy and foreign body mechanics. Flat, circular objects may lie transversely in the relatively narrow cricopharyngeal segment, increasing resistance to spontaneous passage, whereas elongated or irregularly shaped objects may transit the proximal esophagus before lodging more distally. Although this association narrowly missed conventional statistical significance in the univariate analysis, the observed pattern is consistent with evidence suggesting that object morphology and anatomical location influence impaction behavior and subsequent therapeutic decision-making (4,9,10).

The independent association between coin ingestion and proximal esophageal impaction after multivariable adjustment suggests that impaction risk is determined by the interaction between patient-related and object-related factors, rather than by foreign body type alone. Adjustment for age and sex strengthened the association observed in the univariate analysis, suggesting that age may have acted as an important confounding factor. Recent pediatric studies similarly indicate that anatomical maturation and foreign body characteristics jointly influence the location and clinical consequences of esophageal impaction (10,12).

This interpretation is also biomechanically plausible. The cricopharyngeal region is the narrowest physiological section of the esophagus in the pediatric population, and objects with relatively large transverse diameter and flat circular configuration are more likely to lodge at this level. Coins of approximately 20-26 mm may orient perpendicular to the esophageal lumen and therefore face increased resistance to spontaneous passage. Current pediatric literature supports the role of object dimensions, morphology, and anatomical constraints in proximal esophageal impaction.

Object-specific dimensions were not incorporated directly into the multivariable model because measurements were incomplete and heterogeneous across foreign body categories. Within this cohort, coin ingestion may nevertheless serve as a pragmatic surrogate marker for clinically relevant object size. Accordingly, these findings should be regarded as hypothesis-generating rather than as definitive evidence of causality. Prospective multicenter studies incorporating standardized measurements of object dimensions, symptom duration, and time to presentation may improve risk stratification and support the development of clinically applicable prediction models (11,12).

The significant association between the anatomical level of impaction and the extraction technique highlights the importance of tailoring management to the location of the foreign body rather than adopting a uniform therapeutic approach. At our institution, Magill forceps were preferred for objects impacted at the first physiological esophageal narrowing because this technique permits rapid removal under direct laryngoscopic visualization and may avoid more invasive procedures. Conversely, rigid esophagoscopy was preferred for distally impacted objects requiring controlled extraction under direct visualization. Foley catheter extraction represented an effective minimally invasive option for carefully selected patients with recently ingested, blunt, proximally located coins (10,13,14).

Our management outcomes further support an anatomy-based treatment strategy. Selected blunt foreign bodies were intentionally advanced into the stomach when doing so was considered safer than retrieval. Current guidelines recommend balancing the risks of intervention against the likelihood of spontaneous gastrointestinal passage, particularly for blunt objects that have passed beyond the upper esophageal sphincter and do not have high-risk characteristics (1,8).

The low complication rate in our cohort supports the potential safety of individualized management, based on foreign body characteristics and anatomical location, in experienced pediatric centers. Recent studies and guideline-based reviews suggest that overall treatment safety, procedural success, and the need for additional intervention should be considered together when evaluating management outcomes (10,12,15,16). Our findings are consistent with these principles and support the use of structured, guideline-informed management algorithms in routine clinical practice.

We report the epidemiological features of this study and provide an integrated clinical perspective that examines correlations among patient age, foreign body morphology, anatomical impaction patterns, and procedural decision-making. These findings may help clinicians identify children at higher risk of proximal esophageal impaction and select an appropriate extraction technique. Prospective multicenter studies including standardized morphological measurements and relevant temporal variables are required to refine risk stratification and further individualize management.

Study Limitations

This study has several limitations. First, its retrospective, single-center design may limit the generalizability of the findings. Second, some clinically relevant variables, including symptom duration, time from ingestion to presentation, and detailed measurements of all foreign body dimensions, were not consistently available because of the retrospective nature of the study. Third, object-specific dimensions were not incorporated into the

multivariable regression model because of subgroup heterogeneity and incomplete measurements; this omission may have influenced the observed associations. The study reflects the experience of a tertiary pediatric emergency department and therefore may not be directly applicable to institutions with different patient populations or management protocols.

Despite these limitations, this study has several important strengths. It represents one of the largest recently reported single-center pediatric foreign-body ingestion cohorts from Türkiye and provides a comprehensive evaluation of patient demographics, foreign-body characteristics, anatomical impaction patterns, extraction techniques, and clinical outcomes over a five-year period. In addition, the integration of epidemiological findings with biomechanical interpretation and multivariable analysis offers clinically relevant insights that may facilitate individualized management and guide prospective studies.

Conclusion

In conclusion, younger age and coin ingestion were identified as independent predictors of proximal esophageal impaction in children with foreign-body ingestion. These readily available clinical factors may facilitate early risk stratification and support individualized management based on patient age, foreign-body morphology, and anatomical location, rather than on the type of ingested object alone. In our cohort, coins measuring approximately 20-26 mm were more frequently associated with proximal esophageal impaction, suggesting that object diameter may also contribute to impaction risk, particularly in younger children.

An anatomy-based treatment strategy may optimize procedural planning and clinical outcomes. Magill forceps may be considered the preferred first-line technique for foreign bodies impacted at the first physiological esophageal narrowing when adequate visualization is achieved, whereas Foley catheter extraction may serve as a minimally invasive alternative in carefully selected patients. Rigid esophagoscopy remains the gold standard for distally impacted foreign bodies. Overall, applying these anatomy-based principles may improve the selection of extraction techniques, reduce unnecessary invasive procedures, and maintain high procedural success with minimal morbidity. The proposed model demonstrated acceptable discrimination and may support clinical decision-making in pediatric emergency practice. Prospective multicenter studies are warranted to validate these findings and further refine evidence-based management strategies.

Ethics

Ethics Committee Approval: This study was approved by the University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital Clinical Research Ethics Committee (decision no: 301817582, date: 20.01.2026). Ethics approval was obtained retrospectively because the study had a retrospective design and used anonymized patient data.

Informed Consent: Individual informed consent was waived due to the retrospective study design.

Footnotes

Authorship Contributions

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

Conflict of Interest: The authors declared that there were no conflicts of interest.

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

References

1. Ledder O, Furlano R, Papadopoulou A, et al. Foreign body ingestions in children and adolescents: a position paper of the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition (ESPGHAN) endoscopy special interest group. J Pediatr Gastroenterol Nutr. 2026;83:539-54.
2. Stone DE, Lembo BC. Foreign body ingestion and aspiration. Pediatr Rev. 2026;47:352-5.
3. Expert Panel on Pediatric Imaging; Meyers ML, Moore MM, Baker JB, et al. ACR appropriateness criteria® ingested or aspirated foreign body-child. J Am Coll Radiol. 2026;23:732-47.
4. Rogalidou M. Ingestion of foreign bodies and caustic substances in children: a narrative review on clinical evaluation and management update. Clin Exp Pediatr. 2026;69:11-21.
5. Li Z, Yang C, Duan J, et al. Clinical experiences of 109 children with foreign body ingestion: a retrospective study from Kunming, China. Front Pediatr. 2026;14:1771328.
6. Xue N, Wei X, Xu J, Zhu L. High-risk foreign body ingestion in children: clinical features and endoscopic outcomes. Front Pediatr. 2026;14:1832241.
7. Nugud AA, Tzivinikos C, Assa A, et al.; Gastrointestinal Committee of ESPGHAN. Pediatric magnet ingestion, diagnosis, management, and prevention: a European Society for Paediatric Gastroenterology Hepatology and Nutrition (ESPGHAN) position paper. J Pediatr Gastroenterol Nutr. 2023;76:523-32.
8. Kramer RE, Lerner DG, Lin T, et al.; North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition Endoscopy Committee. Management of ingested foreign bodies in children: a clinical report of the NASPGHAN Endoscopy Committee. J Pediatr Gastroenterol Nutr. 2015;60:562-74.
9. Di Siena A, Melli P, Rodofile F, et al. Foreign body ingestion in children: a 16-year experience. Pediatr Gastroenterol Hepatol Nutr. 2025;28:215-23.
10. Pereira LDM, Barreira MA, de Saboia Mont'Alverne TN, et al. Endoscopic techniques and factors for complications in pediatric

esophageal foreign body removal. World J Gastrointest Endosc. 2025;17:106352.

11. Manfredi MA, Alvarez RP, Arai K, et al. Global insights on the diagnosis, management, and prevention of pediatric ingestions: a report from the FISPGHAN expert panel. JPGN Rep. 2025;6:274-87.
12. Yeon HJ, Lee SM, Ihn K, Oh JT, Han SJ, Ho IG. Pediatric foreign body ingestion: analysis of patient characteristics and surgical treatment. Children (Basel). 2025;12:1355.
13. Erginel B, Kaba M, Karadag CA, Yildiz A, Demir M, Sever N. Foley catheter technique for the extraction of coins lodged in the upper esophagus of children. BMC Pediatr. 2023;23:605.
14. Kasakanga K, Groenewald A, Mabitsela ME, Tshifularo N. A retrospective review of the management of impacted coin ingested in children. Afr J Paediatr Surg. 2025;22:32-7.
15. Hashimi B, Shaffer AD, McCoy JL, Chi DH, Padia R. Resource utilization and risk factors for esophageal injury in pediatric esophageal foreign bodies. Laryngoscope. 2024;134:4774-82.
16. Demiroren K. Management of gastrointestinal foreign bodies with brief review of the guidelines. Pediatr Gastroenterol Hepatol Nutr. 2023;26:1-14.

Original Article

DOI: 10.4274/haseki.galenos.2026.27247
Med Bull Haseki 2026;64(5):435-441

Predictors of Stone-free Status Following Ultrathin Semi-rigid Ureteroscopy for Mid and Distal Ureteral Stones in Children

Ayberk Iplikci[1], Muhammet Cicek[2], Ozgur Arikan[3], Mammad Khalilov[3], Mesrur Selcuk Silay[4], Asif Yildirim[3]

[1]*Boyabat 75th Year State Hospital, Clinic of Urology, Sinop, Türkiye*
[2]*Malatya Training and Research Hospital, Clinic of Urology, Malatya, Türkiye*
[3]*Istanbul Medeniyet University Faculty of Medicine, Department of Urology, Istanbul, Türkiye*
[4]*Medipol Acibadem Hospital, Clinic of Urology, Istanbul, Türkiye*

Abstract

Aim: Pediatric ureteral stones present a clinical challenge because of anatomical limitations and the need for minimally invasive treatment approaches. The aim of this study was to assess outcomes of semi-rigid ureteroscopy (URS) in pediatric patients with mid- and distal-ureteral stones, with particular focus on predictors of postoperative stone-free (SF) status.

Methods: This single-center retrospective observational study included pediatric patients who underwent 4.5- and 6.5-Fr URS for mid- and distal-ureteral stones smaller than 2 cm between December 2016 and December 2024. Demographic characteristics, stone features, and perioperative outcomes were recorded. Participants were divided into two groups defined by the presence (n=17) or absence (n=52) of residual stones. Univariate and multivariate regression analyses were performed to identify factors associated with postoperative SF status.

Results: The study included 69 children (median age, 9 years; median stone size, 8 mm). Postoperative imaging revealed an SF rate of 75.4%. A history of surgery was less common in the SF group ($p<0.001$). The two groups differed in stone burden ($p<0.001$) and number of stones ($p<0.001$). Four (5.8%) patients had postoperative febrile urinary tract infections. Multivariate regression analysis revealed that stone burden (>10 mm vs. ≤10 mm) was the only predictor for residual fragments after URS for mid and distal ureteral stones ($p=0.048$).

Conclusion: Ureteroscopy is a safe, effective, and applicable minimally invasive treatment modality with low complication rates for ureteral stones in children. However, stone burden should be considered during patient selection and preoperative counseling, as it independently predicts postoperative SF status after ultrathin pediatric URS.

Keywords: Pediatric urolithiasis, ureteral stone, ureteroscopy, stone burden, stone-free status

Introduction

Urolithiasis represents a significant cause of morbidity across both adult and pediatric populations, exhibiting a globally escalating prevalence (1,2). Pediatric urolithiasis is characterized by a high recurrence rate, attributable to its frequent association with underlying metabolic, anatomical, and infectious etiologies (3). Due to the high recurrence rate of pediatric urolithiasis and the likelihood of multiple surgical interventions, minimally invasive treatment is essential (4). Achieving stone-free (SF) status remains a paramount objective in clinical management (5). Treatment strategies are determined by a confluence of factors, including stone location, stone size, and the patient's unique urinary tract anatomy.

Corresponding Author: Ayberk Iplikci, MD, Boyabat 75th Year State Hospital, Clinic of Urology, Sinop, Türkiye
E-mail: ayberkiplikci@gmail.com **ORCID:** orcid.org/0000-0002-5822-7799
Received: 23.12.2025 **Accepted:** 16.07.2026 **Publication Date:** 17.11.2026

Cite this article as: Iplikci A, Cicek M, Arikan O, Khalilov M, Silay MS, Yildirim A. Predictors of stone-free status following ultrathin semi-rigid ureteroscopy for mid and distal ureteral stones in children. Med Bull Haseki. 2026;64(5):435-441

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

Ureteral stones account for approximately 7% of all urinary calculi in children (6). The management options for children presenting with ureteral stones include a trial of spontaneous passage, medical expulsive therapy (MET), semi-rigid ureteroscopy (URS), and extracorporeal shock wave lithotripsy. Pediatric ureteral stones below 5 mm frequently pass spontaneously, while larger stones usually require intervention. Shock wave lithotripsy is associated with limitations, notably reduced efficacy, the necessity for serial treatments and subsequent anesthetic events, and a resultant elevation in cumulative radiation dosage (7). Ureteroscopy is often prioritized due to its expeditious achievement of stone clearance and minimization of patient discomfort (8). Because URS is primarily used for mid and distal ureteral stones and treatment outcomes may vary according to stone location, we focused on this relatively homogeneous patient population. Excluding proximal ureteral stones allowed a more accurate evaluation of factors associated with postoperative SF status.

However, the existing literature on URS in pediatric patients is limited by several factors, including the lack of standardized surgical protocols, variability in instrument diameters, and heterogeneity in intracorporeal lithotripsy modalities. We hypothesized that ultrathin, semi-rigid URS would provide high SF rates and low complication rates in children with mid- and distal ureteral stones and that stone-related factors would be associated with postoperative SF status. In this context, we assessed the outcomes of 4.5/6.5 Fr (ultrathin) URS in children with mid- and distal-ureteral stones, focusing on predictors of postoperative SF status.

Materials and Methods

Compliance with Ethical Standards

A retrospective analysis was conducted on 69 pediatric patients who underwent unilateral URS between December 2016 and December 2024 for the treatment of mid- and distal ureteral stones measuring less than 2 cm. This single-center study was conducted in accordance with the Declaration of Helsinki. Ethical approval for the study was obtained from the Goztepe Prof. Dr. Suleyman Yalcin City Hospital Non-Interventional Clinical Research Ethics Committee (approval no.: 2025/0248, date: 23.10.2025).

Study Design

All patients underwent preoperative low-dose non-contrast computed tomography (NCCT). Patients with abnormal renal anatomy, proximal ureteral stones, pushback during surgery, ureteral stenosis, or incomplete medical records were excluded from the study. Given the established consensus that ureteral stones exceeding 5 mm typically necessitate endourological intervention, we included patients with such stones. Patients with preoperative stents were not included in the study. The patient selection process and the study cohort are summarized in the flow diagram (Figure 1).

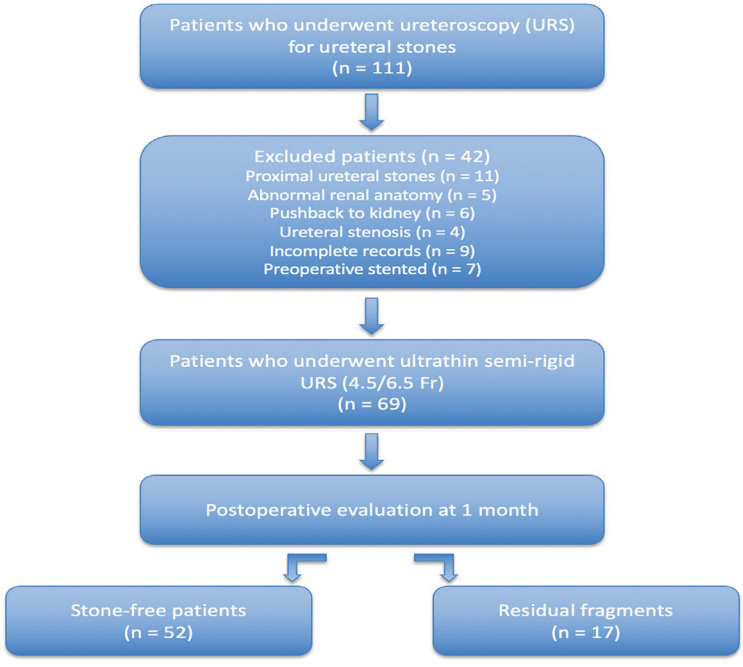


Figure 1. Study flow diagram showing patient selection and study cohort

Demographic data, degree of hydronephrosis, stone characteristics (laterality, number, location, size, and density), operative time, and perioperative and postoperative complications were recorded. Stone size was defined as the longest diameter of the stone, regardless of the NCCT axes. Stone density (Hounsfield unit) was quantified by manual delineation of the region of interest, which precisely conformed to the stone margins. In the case of multiple stones, stone burden was considered the sum of the stones' longest diameters. The Society for Fetal Urology grading system was used to assess hydronephrosis severity on NCCT.

Operative Details

A negative urine culture was confirmed in all participants prior to the surgical intervention. All surgical procedures were conducted under general anesthesia. Prophylactic antibiotics were administered at a dose calculated according to guidelines. Case-specific surgical teams were assembled due to the long study period. However, all procedures were performed by urologists experienced in pediatric stone surgery, each of whom had performed at least 20 pediatric stone procedures before participating in the study. The same surgical principles and operative technique were applied throughout the study period. A 4.5/6.5 Fr semi-rigid ureteroscope (ultrathin; Richard Wolf GmbH, Knittlingen, Germany) served as the endoscopic instrument. Prior to the introduction of a safety guidewire (0.035-inch flexible) to the kidney, an examination of the urethra and bladder was performed. Fragmentation of the calculus was accomplished using a 272 µm Holmium:YAG laser fiber (Quanta System, Milan, Italy), with laser energy delivered in the range of 6 to 10 Watts. The dimensions of the stone fragments were assessed relative to the guidewire; fragments with a diameter equal to or smaller than that of the guidewire were left *in situ*. Stone fragments were irrigated with saline solution from the ureter to the bladder to promote spontaneous expulsion. After lithotripsy, a Foley catheter was inserted into all participants. The necessity and dimensions of double-J stent insertion were determined on an individualized basis for each patient. The urethral catheter was removed on the first postoperative day, and patients were subsequently discharged. In selected cases, extraction strings were used on the first postoperative day for stent removal.

Follow-up Procedure

At the first postoperative month, SF status was assessed by kidney-ureter-bladder (KUB) radiography and ultrasonography or by NCCT in cases of suspected residual calculus. To mitigate radiation exposure, NCCT was not routinely performed in all patients. Stone-free status was defined as the patient being completely SF or having residual fragments measuring less than 2 mm on radiological imaging. Double-J stents were removed within two weeks of the procedure in all patients, unless clinical indications necessitated extended stent retention. The primary endpoint of the study was postoperative SF status at the first-month follow-up.

Statistical Analysis

The data was analyzed using SPSS version 26 (Armonk, NY: IBM Corp., USA). Categorical variables in the study were presented with frequency (n) and percentage. For quantitative data, the Shapiro-Wilk test was used to assess the normality of the distribution. Quantitative variables with normal distribution were specified as mean ± standard deviation, and variables without normal distribution were presented as median (range) for descriptive statistics in the study. An independent samples t-test was used to compare the means. When the variables between groups were not normally distributed, the Mann-Whitney U test was used. For the analysis of categorical variables, the chi-squared test and Fisher's exact test were used to compare the two groups. Univariate and multivariate logistic regression analyses were performed to determine the factors associated with residual fragments in the first month after surgery. Variables considered clinically relevant and those associated with residual fragments in the univariate analysis were included in the multivariable logistic regression model. Statistical significance was set at a p-value <0.05.

Results

Sixty-nine patients, comprising 46 boys and 23 girls, underwent URS. The median age was 9 years (range 1-17) and the median stone size was 8 mm (range 5-20). Postoperative imaging revealed an SF rate of 75.4% (n=52). Participants were divided into two groups based on SF status, and comparative analyses were performed. Statistical analysis revealed no significant differences between the groups regarding gender (p=0.429), age (p=0.710), family history (p=0.364), preoperative hydronephrosis (p=0.530), stone laterality (p=0.720), maximum stone size (p=0.187), stone density (p=0.928), or operation time (p=0.195). Surgical history was less common in the SF group (p<0.001). A difference between the two groups was observed for stone burden (p<0.001) and number of stones (p<0.001) (Table 1).

The size of the ureteral stent was determined on a case-by-case basis; however, in 14 patients (20.3%), stent placement was not indicated due to the absence of clinical need. Within 24 hours of the procedure, 66 patients (95.6%) were discharged. Four patients (5.8%) had postoperative febrile urinary tract infection. Stone analysis was not feasible in 71% of participants (Table 2). An auxiliary URS procedure was required for 10 (14.5%) children who were not SF after the initial surgery.

Multivariate regression analysis revealed stone burden (>10 mm vs. ≤10 mm) was the only predictor for residual fragments after URS for mid- and distal ureteral stones (p=0.048) (Table 3).

Table 1. Demographic and clinical characteristics of the groups			
	Stone-free (n=52)	**Non-stone-free (n=17)**	**p-value**
Gender, n (%) **Girl** **Boy**	 16 (30.8) 36 (69.2)	 7 (41.2) 10 (58.8)	0.429[a]
Age (years), median (min-max)	10 (1-17)	7 (1-17)	0.710[b]
Family history, n (%)	24 (46.1)	10 (58.8)	0.364[a]
Surgical history, n (%)	11 (21.1)	14 (82.3)	**<0.001[a]**
Preoperative hydronephrosis, n (%) **None** **Grade 1** **Grade 2** **Grade 3-4**	 8 (15.4) 21 (40.4) 21 (40.4) 2 (3.8)	 2 (11.8) 4 (23.5) 10 (58.8) 1 (5.9)	0.530[a]
Laterality, n (%) **Right** **Left**	 24 (46.1) 28 (53.8)	 7 (41.2) 10 (58.8)	0.720[a]
Number of stones, n (%) **Single** **Multiple**	 48 (92.3) 4 (7.7)	 8 (47) 9 (53)	**<0.001[a]**
Maximum stone size (mm), median (min-max)	7 (5-20)	8 (5-20)	0.187[b]
Stone burden (mm), median (min-max)	8 (5-20)	14 (5-28)	**<0.001[b]**
Stone density (Hounsfield unit), median (min-max)	762 (350-1735)	900 (450-1700)	0.928[b]
Operation time (minutes), median (min-max)	40 (15-90)	45 (20-70)	0.195[b]

[a]Pearson's chi-squared test, [b]Mann-Whitney U test

Table 2. Postoperative clinical characteristics of participants	
	Participants (n=69)
Ureteral stent placement, n (%) **None** **3 Fr Double-J stent** **4 Fr Double-J stent** **4.7 Fr Double-J stent**	 14 (20.3) 16 (23.2) 13 (18.8) 26 (37.7)
Postoperative febrile UTI, n (%)	4 (5.8)
Length of hospital stay, n (%) **1 day** **>1 day**	 66 (95.6) 3 (4.3)
Stone analysis, n (%) **Not performed** **Calcium oxalate** **Cystine**	 49 (71) 7 (10.1) 13 (18.8)

UTI: Urinary tract infection

Table 3. Univariate and multivariate analysis: factors affecting residual fragments after URS				
	Univariate analysis		**Multivariate analysis**	
	OR (95% CI)	**p-value**	**OR (95% CI)**	**p-value**
Age (>10 years vs. ≤10 years)	0.809 (0.308-2.126)	0.668		
Surgical history (yes vs. no)	8.213 (2.360-28.580)	**0.001**	3.646 (0.899-14.782)	0.070
Number of stones (multiple vs. single)	4.846 (1.870-12.560)	**0.001**	1.237 (0.414-3.702)	0.703
Stone burden (>10 mm vs. ≤10 mm)	9.333 (2.682-32.477)	**<0.001**	4.390 (0.979-19.691)	**0.048**
Stone density (>800 HU vs. ≤800 HU)	1.887 (0.698-5.103)	0.211		

Bold values indicate statistically significant p-values (<0.05)
OR: Odds ratio, CI: Confidence interval, URS: Ureteroscopy, HU: Hounsfield unit

Discussion

Because the propensity for stone recurrence is higher in pediatric patients than in adults, the potential need for future surgical intervention should be considered, and minimally invasive techniques should be the initial therapeutic approach. The clinical management of pediatric urolithiasis is challenged by the inherent fragility and narrow diameter of the collecting system (9). In spite of the challenges presented, technological innovations, specifically the miniaturization of endoscopic equipment, have improved the safety and efficacy of minimally invasive approaches in clinical practice (10,11). The development of ultrathin URS was undertaken to achieve a reduction in complication rates (12). However, a recent survey showed that the availability of ultrathin URS is limited, with only 24% of 74 EAU centers having this equipment (13). In this study, all cases were treated using our routinely used ultrathin URS, which has been available for pediatric stone procedures for many years. In our study, we used the Holmium:YAG laser for lithotripsy. While pneumatic lithotripters were previously common in pediatric URS, they are now rarely used due to higher complication rates and lower SF rates (9,14,15).

The primary objective of URS for the management of ureteral calculi is the attainment of complete stone clearance. Published reports indicate URS success rates ranging from 75% to 100%, with superior outcomes observed for distal ureteral stones (7,10,16-18). Our SF rate (75.4%) is consistent with the existing literature. SF status was defined as the absence of fragments or the presence of fragments <2 mm, which is a stricter definition than that used in several previous pediatric urolithiasis studies (19,20). Our study cohort encompassed patients with a single ureteral calculus or multiple ureteral calculi. Consequently, we determined that stone burden affected SF status after ultrathin URS. The literature contains extensive evidence of a negative correlation between stone size and postoperative SF rates in children; however, studies evaluating the impact of stone burden on procedural success are limited (15,21,22). Recent studies have identified ureteral wall thickness and ureteral diameter as potential predictors of SF status after URS (23). As these parameters were not routinely recorded in our retrospective dataset, they could not be evaluated in the present study. Prospective studies may further clarify their role in predicting procedural success.

Ureteroscopy provides effective treatment with low morbidity and high patient satisfaction. Postoperative minor complications, including mucosal injury, hematuria, fever, and colic pain, may occur, with reported rates not exceeding 20% (24,25). Dogan et al. (24) reported an 8.4% complication rate in a retrospective, multi-institutional pediatric URS study, with operative time identified as the only factor affecting complication rates. In this study, laser use was limited to one out of five patients, and instruments larger than 8 Fr were used in more than half of the patients (24). Our study demonstrated a low complication rate: transient fever was the only observed adverse event, affecting four patients (5.8%). A 95.6% next-day discharge rate was achieved, attributable to a remarkably low complication profile. The use of small-caliber URS and the limited duration of the procedures are hypothesized to contribute to this favorable outcome. In our study, pushback to the kidney was not observed; this finding is attributable to the absence of proximal ureteral calculi. Furthermore, major complications such as perforation, ureteral rupture, or sepsis, which have been documented in the literature, did not occur in our series (26).

In our study, active dilation of the ureteral orifice was intentionally omitted in all patients due to the potential for complications, including ureteral stenosis and vesicoureteral reflux (27). Two recent investigations have reported access failure rates of 16.6% and 20% when employing 6 Fr and 7.5 Fr semi-rigid ureteroscopes in pediatric patients, respectively (22,28). In concordance with findings reported in previous studies, ultrathin URS facilitated successful ureteral access in all patients, notwithstanding the notably younger age distribution of our cohort (29,30). Postoperative stenting remains a subject of debate. However, it is believed to promote enhanced healing and resolution of local edema, while also mitigating postoperative pain and the risk of infectious complications arising from obstruction (26). Ureteral stenting after pediatric URS is reported in 31-100% of cases (10). However, stent removal necessitates an additional procedure under general anesthesia. Given the relatively young age of our patient cohort and the potential for postoperative renal colic secondary to mucosal edema, temporary catheter placement was performed in 55 of 69 patients. In addition, despite the demonstrated efficacy of MET in children, its off-label status and potential safety concerns precluded its use in our study (8,31,32).

Study Limitations

Our study has limitations, including its retrospective design. Additionally, we primarily used KUB radiography and ultrasonography instead of NCCT to assess SF status to reduce radiation exposure in children; this choice may have influenced the reported SF rate. The lower SF rate observed in our cohort, compared with some reports in the literature, may also be explained by the inclusion of patients with multiple stones and a higher overall stone burden. Other significant limitations include the limited number of patients and the wide age range. In addition, the relatively small number of residual stone events (n=17) may have

increased the risk of overfitting in the multivariable logistic regression model. Therefore, the identified predictors should be interpreted with caution and validated in larger prospective cohorts. Stone analysis was not available for all patients because of our predominant use of the dusting technique. Furthermore, due to limitations inherent in the retrospective dataset, a subgroup analysis comparing SF outcomes between mid and distal ureteral stones could not be performed. While the same surgical technique was consistently employed by all surgeons within this cohort, the possibility of inter-surgeon differences affecting both success and complication rates remains. This factor should be acknowledged as a limitation that influences the demonstrative capacity of the data. Despite these limitations, this study makes a significant contribution to the literature as one of the largest series validating the safety and efficacy of ultrathin URS for pediatric ureteral stone management.

Conclusion

Ureteroscopy is a safe, effective, and feasible minimally invasive treatment modality with low complication rates for ureteral stones in children. Our study demonstrates that URS in children yields favorable outcomes, characterized by short hospital stays and high SF rates. The success of the procedure is significantly affected by the stone burden. Therefore, stone burden should be considered during patient selection and preoperative counseling because it independently predicts postoperative SF status after ultrathin pediatric URS.

Ethics

Ethics Committee Approval: The ethics committee approved the commencement of the present study Goztepe Prof. Dr. Suleyman Yalcin City Hospital Non-Interventional Clinical Research Ethics Committee (approval no.: 2025/0248, date: 23.10.2025). All authors fully complied with the Declaration of Helsinki during the course of the study.

Informed Consent: The ethics committee waived the requirement for informed consent because the study was retrospective.

Authorship Contributions

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

Conflict of Interest: No conflicts of interest were declared by the authors.

Financial Disclosure: This study received no financial support.

References

1. Sorokin I, Mamoulakis C, Miyazawa K, Rodgers A, Talati J, Lotan Y. Epidemiology of stone disease across the world. World J Urol. 2017;35:1301-20.
2. Utangaç MM, Dede O. Evaluation of the largest series of ultramini percutaneous nephrolithotomy in preschool children: 10-year experience with 711 units of kidney stones. J Clin Med. 2025;14:3355.
3. Gajengi AK, Wagaskar VG, Tanwar HV, Mhaske S, Patwardhan SK. Metabolic evaluation in paediatric urolithiasis: a 4-year open prospective study. J Clin Diagn Res. 2016;10:PC04-6.
4. Baştuğ F, Gündüz Z, Tülpar S, Poyrazoğlu H, Düşünsel R. Urolithiasis in infants: evaluation of risk factors. World J Urol. 2013;31:1117-22.
5. Chew BH, Brotherhood HL, Sur RL, et al. Natural history, complications and re-intervention rates of asymptomatic residual stone fragments after ureteroscopy: a report from the EDGE research consortium. J Urol. 2016;195:982-6.
6. Raza A, Smith G, Moussa SA, Tolley DA. Ureteroscopy in the management of pediatric urinary tract calculi. J Endourol. 2005;19:151-8.
7. Smaldone MC, Cannon GM, Wu H-Y, et al. Is ureteroscopy first line treatment for pediatric stone disease?. J Urol. 2007;178:2128-31.
8. Radmayr C, Bogaert G, Bujons A et al. EAU Guidelines on Paediatric Urology. 2026. Available from: https://d56bochluxqnz.cloudfront.net/documents/full-guideline/EAU-Guidelines-on-Paediatric-Urology-2026.pdf
9. Yucel S, Akın Y, Kol A, Danışman A, Guntekin E. Experience on semirigid ureteroscopy and pneumatic lithotripsy in children at a single center. World J Urol. 2011;29:719-23.
10. Ishii H, Griffin S, Somani BK. Ureteroscopy for stone disease in the paediatric population: a systematic review. BJU Int. 2015;115:867-73.
11. Onal B, Citgez S, Tansu N, et al. What changed in the management of pediatric stones after the introduction of minimally invasive procedures? A single-center experience over 24 years. J Pediatr Urol. 2013;9:910-4.
12. Herndon CD, Viamonte L, Joseph DB. Ureteroscopy in children: is there a need for ureteral dilation and postoperative stenting?. J Pediatr Urol. 2006;2:290-3.
13. Bañuelos Marco B, Haid B, Radford A, et al. Guideline adherence of paediatric urolithiasis: an EAU Members' Survey and expert panel roundtable discussion. Children (Basel). 2022;9:504.
14. Irsayanto D, Setiawan MR, Salsabila S, et al. Efficacy and safety of pneumatic lithotripsy with laser lithotripsy in the treatment of ureteral stones <20 millimeters in children: a systematic review and meta-analysis. J Ayub Med Coll Abbottabad. 2024;36:201-8.
15. Turunc T, Kuzgunbay B, Gul U, et al. Factors affecting the success of ureteroscopy in management of ureteral stone diseases in children. J Endourol. 2010;24:1273-7.

16. Utanğaç MM, Sancaktutar AA, Tepeler A. Micro-ureteroscopy for the treatment of distal ureteral calculi in children. J Pediatr Surg. 2017;52:512-6.

17. Dogan HS, Tekgul S, Akdogan B, Keskin MS, Sahin A. Use of the Holmium:YAG laser for ureterolithotripsy in children. BJU Int. 2004;94:131-3.

18. Sinha M, Pietropaolo A, Quiroz Madarriaga Y, et al. Outcomes of ureteroscopy for management of stone disease in early and late childhood over a 15-year period. Ther Adv Urol. 2022;14:17562872221141775.

19. Amer ML, Elkholefy ME, Dawoud H, Gaber M, El-Gamal OM, Gameel T. Safety and efficacy of Silodosin as medical expulsive therapy after shock wave lithotripsy in paediatric patients with renal stones. Urolithiasis. 2025;53:95.

20. Ahmed A, Farhan B, Alam-Eldin A, et al. A prospective randomized trial comparing dusting and fragmentation techniques using Holmium:YAG laser for pediatric ureteral stones. Can J Urol. 2025;32:483-90.

21. Jaidane M, Hidoussi A, Slama A, Hmida W, Ben Sorba N, Mosbah F. Factors affecting the outcome of ureteroscopy in the management of ureteral stones in children. Pediatr Surg Int. 2010;26:501-4.

22. Kucukdurmaz F, Efe E, Sahinkanat T, Amasyalı AS, Resim S. Ureteroscopy with Holmium:YAG laser lithotripsy for ureteral stones in preschool children: analysis of the factors affecting the complications and success. Urology. 2018;111:162-7.

23. Selvi I, Baydilli N, Tokmak TT, Akinsal EC, Basar H. CT-related parameters and Framingham score as predictors of spontaneous passage of ureteral stones ≤ 10 mm: results from a prospective, observational, multicenter study. Urolithiasis. 2021;49:227-37.

24. Dogan HS, Onal B, Satar N, et al. Factors affecting complication rates of ureteroscopic lithotripsy in children: results of multi-institutional retrospective analysis by Pediatric Stone Disease Study Group of Turkish Pediatric Urology Society. J Urol. 2011;186:1035-40.

25. Topaktas R, Aydin C, Altin S, Akkoc A, Aydın ZB, Urkmez A. The efficacy of ultra-thin semi-rigid ureteroscopy with holmium laser lithotripsy in pediatric ureteral stones: a single-center experience. Cureus. 2019;11:e5496.

26. Tiryaki T, Azili MN, Özmert S. Ureteroscopy for treatment of ureteral stones in children: factors influencing the outcome. Urology. 2013;81:1047-51.

27. Hubert KC, Palmer JS. Passive dilation by ureteral stenting before ureteroscopy: eliminating the need for active dilation. J Urol. 2005;174:1079-80; discussion 1080.

28. AbdelRazek M, Fathi A, Mohamed M, Abdel-Kader MS. Primary versus deferred ureteroscopy for the management of obstructive anuria secondary to ureteric urolithiasis in children: a prospective randomized study. Urolithiasis. 2022;51:6.

29. Atar M, Sancaktutar AA, Penbegul N, et al. Comparison of a 4.5 F semi-rigid ureteroscope with a 7.5 F rigid ureteroscope in the treatment of ureteral stones in preschool-age children. Urol Res. 2012;40:733-8.

30. Kocaoglu C, Ozkan KU. The effectiveness of 4.5F ultrathin semirigid ureteroscope in the management of ureteral stones in prepubertal children: is there a need for any ureteral dilatation?. Urology. 2014;84:202-5.

31. Mokhless I, Zahran AR, Youssif M, Fahmy A. Tamsulosin for the management of distal ureteral stones in children: a prospective randomized study. J Pediatr Urol. 2012;8:544-8.

32. Elgalaly H, Eliwa A, Seleem M, et al. Silodosin in the treatment of distal ureteric stones in children: a prospective, randomised, placebo-controlled study. Arab J Urol. 2017;15:194-8.

Original Article

DOI: 10.4274/haseki.galenos.2026.93585
Med Bull Haseki 2026;64(5):442-448

Guideline Concordance, Information Quality, and Readability of Large Language Model Responses to Female Urethral Stricture Questions: A Comparative Cross-sectional Study

Cagri Sevik, Metin Savun, Suleyman Sahin, Yigit Can Filtekin, Halil Lutfi Canat

University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Urology, Istanbul, Türkiye

Abstract

Aim: Large language models (LLMs) are increasingly employed to acquire medical information. However, their reliability in addressing rare and technically intricate conditions, such as female urethral stricture disease (FUSD), remains uncertain. This study evaluated the guideline concordance, information quality, reliability, and readability of responses generated by three widely accessible LLMs in response to clinical questions related to FUSD.

Methods: This comparative cross-sectional study was conducted in January 2026. Thirty standardized questions across five clinical domains were developed based on the 2025 European Association of Urology Guidelines for Urethral Stricture Disease and refined through expert consensus. These questions were submitted to ChatGPT 5.0, Gemini 3.0, and Microsoft Copilot (Basic) under identical conditions. Responses were independently evaluated by three blinded board-certified urologists using a 5-point expert-rated accuracy scale. Information reliability and quality were assessed using modified DISCERN and ensuring quality information for patients (EQIP) independently by the blinded reviewers, while readability was evaluated using flesch reading ease score (FRES) and flesch-kincaid grade level (FKGL). Between-model comparisons were performed using the Friedman test, followed by Wilcoxon signed-rank tests with Bonferroni correction.

Results: Gemini achieved the highest expert-rated accuracy (4.80±0.41) and modified DISCERN score (19.0±1.7) among the three models. The Copilot achieved the highest EQIP (89.60±4.01) and FRES (42.23±3.89) scores and the lowest FKGL (12.67±1.04). Significant differences between the models were identified across all evaluated metrics. Gemini had significantly higher modified DISCERN scores than ChatGPT and Copilot (both $p<0.001$), whereas ChatGPT and Copilot did not differ significantly.

Conclusion: The evaluated LLMs demonstrated distinct performance profiles rather than uniform superiority. Gemini showed the highest expert-rated accuracy and information reliability, whereas Copilot provided the most favorable patient-oriented information quality and readability. Large language model-generated information on FUSD should be verified against current clinical guidance before clinical or patient education.

Keywords: Artificial intelligence, health literacy, large language models, patient education as topic, urethral stricture

Introduction

Female urethral stricture disease (FUSD) is an uncommon yet clinically significant condition that contributes to lower urinary tract symptoms in women. Its clinical presentation is often non-specific, including reduced urinary flow, incomplete bladder emptying, frequency, urgency, recurrent urinary tract infections, and urinary incontinence, which may contribute to delayed diagnosis (1). The diagnostic evaluation of FUSD requires careful clinical assessment and may include uroflowmetry, post-void residual measurement, cystourethroscopy, and radiographic or urodynamic evaluation. Treatment options range from urethral dilation and intermittent

Corresponding Author: Cagri Sevik, MD, University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Urology, Istanbul, Türkiye
E-mail: cagri.sevik@hotmail.com **ORCID:** orcid.org/0000-0003-0288-704X
Received: 16.07.2026 **Accepted:** 30.10.2026 **Publication Date:** 17.11.2026

Cite this article as: Sevik C, Savun M, Sahin S, Filtekin YC, Canat HL. Guideline concordance, information quality, and readability of large language model responses to female urethral stricture questions: a comparative cross-sectional study. Med Bull Haseki. 2026;64(5):442-448

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

self-dilation to reconstructive urethroplasty, with the choice of intervention depending on previous treatments, recurrence, stricture characteristics, and patient-related factors (1,2). Despite increasing clinical awareness, variability in definitions, diagnostic criteria, treatment strategies, and outcome reporting remains an important challenge in FUSD.

Large language models (LLMs) are increasingly used by both patients and healthcare professionals to obtain medical information and support patient education (3). Recent studies in urology have demonstrated that LLM-generated responses may provide generally accurate and informative content; however, important differences have been reported among models in terms of information quality, reliability, readability, and guideline concordance (4,5). In reconstructive urology, ChatGPT has shown potential for answering common patient questions, although limitations in readability and information quality remain (6,7). Similarly, recent comparative studies evaluating artificial intelligence (AI) chatbots in other functional urological conditions have demonstrated substantial model-dependent variation in both the quality and accessibility of generated health information (8,9). These findings suggest that the performance observed in common urological conditions cannot necessarily be extrapolated to less prevalent and technically complex disorders.

However, evidence regarding the performance of contemporary LLMs in FUSD remains limited. This is particularly relevant because FUSD combines diagnostic uncertainty, multiple treatment pathways, technically demanding reconstructive procedures, and the need for clear patient counseling, creating distinct challenges for AI-generated medical information. We hypothesized that the evaluated LLMs would differ significantly in terms of guideline concordance, information quality, and readability when responding to standardized questions related to FUSD. Therefore, the present study aimed to compare the responses generated by three widely available LLMs across the major clinical domains of FUSD, including diagnosis, treatment, surgical management, follow-up, and patient counseling, using expert assessments, validated information-quality instruments, and established readability metrics.

Materials and Methods

Study Design

This comparative cross-sectional study was conducted in January 2026 to evaluate the performance of the three LLMs in responding to standardized clinical questions regarding FUSD. The evaluated models were ChatGPT 5.0 (OpenAI), Gemini 3.0 (Google), and Copilot Basic (Microsoft). This study exclusively evaluated responses generated by publicly accessible AI platforms and did not involve human participants, animals, biological materials, patient records, or personally identifiable data. Therefore, ethics committee approval and informed consent were not required for this study.

Questionnaire Development

A set of 30 standardized clinical questions was used to assess the performance of the LLMs across major aspects of FUSD management. The questions were developed based on the 2025 European Association of Urology Guidelines for Urethral Stricture Disease (1). They were categorized into five predefined clinical domains: diagnostic evaluation (eight questions), treatment and initial management (eight questions), surgical considerations (six questions), follow-up and recurrence (four questions), and patient counseling (four questions). For transparency and reproducibility, the complete set of questions and their corresponding guidelines or reference sources are provided in Supplementary Table 1.

Large Language Model Testing Conditions

Freely accessible versions of all three LLMs were used through their official web interfaces. An identical, standardized prompt structure and the same clinical questions were used for all models. Each question was submitted as a separate conversation. The chat history and memory functions were disabled during testing to minimize contextual carryover between queries. No follow-up prompts were used, and the responses were left unchanged. Therefore, only the first response generated for each question was included in the analyses. The generated responses were subsequently collected for a blinded expert evaluation.

Expert Evaluation

Three board-certified urologists independently evaluated the responses. Before the evaluation, all model identifiers were removed, and the reviewers were blinded to the identity of the LLM that generated the responses. Each reviewer independently assessed the responses without discussion or consensus scoring during the initial evaluation.

Outcome Measures

Expert-rated accuracy and guideline concordance were assessed using a 5-point Likert scale. Responses were evaluated against the guidelines or reference standards corresponding to each clinical question, as detailed in Supplementary Table 1. The scores were interpreted as follows: 1, completely incorrect or guideline-discordant; 2, mostly incorrect with major errors or clinically relevant omissions; 3, partially correct with relevant omissions; 4, mostly correct and guideline-concordant with only

minor omissions; and 5, fully correct, complete, and guideline-concordant. For each response, the final expert-rated accuracy score used in the statistical analysis was calculated as the mean of the scores assigned by all three reviewers.

The reliability and quality of the generated health information were additionally assessed using the modified DISCERN instrument (comprising 5 core items scored on a 1-to-5 scale, yielding a total score ranging from 5 to 25) and the ensuring quality information for patients (EQIP) tool (consisting of 36 items evaluated to calculate a standardized percentage score). These tools were independently applied and scored by the blinded expert reviewers to ensure objective and standardized evaluation. Higher scores indicated higher reliability and higher information quality, according to the respective instrument. The scoring procedure and score range for each instrument were applied consistently across all model responses. The readability of each generated response was assessed using the flesch reading ease score (FRES) and flesch-kincaid grade level (FKGL). Higher FRES values indicated easier readability, whereas higher FKGL values indicated a higher educational grade level required to understand the text and therefore greater textual complexity.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics, version 27 (IBM Corp., Armonk, NY, USA). Continuous variables are presented as mean ± standard deviation. For each response, the scores assigned by the three reviewers were averaged before the between-model comparisons were performed. Inter-rater agreement among the three reviewers was assessed using the intraclass correlation coefficient (ICC) based on a two-way random-effects model with absolute agreement.

Because the same set of 30 questions was evaluated across all three LLMs, the observations constituted paired repeated measurements. Overall differences among ChatGPT, Gemini, and Copilot were therefore assessed using the Friedman test. When a significant overall difference was identified, pairwise comparisons were performed using the Wilcoxon signed-rank test. To account for the three pairwise comparisons, statistical significance for post-hoc analyses was evaluated using a Bonferroni-adjusted threshold of $p<0.0167$. Kendall's W was calculated as an effect-size measure for the Friedman test. A two-sided p-value of <0.05 was considered statistically significant for the overall analyses.

ChatGPT 5.0, Gemini 3.0, and Microsoft Copilot served as subjects in this study. Generative AI was additionally used during the manuscript revision to assist with the language, clarity, and structural editing. All AI-assisted content was critically reviewed and approved by the authors, who retained full responsibility for the scientific content and conclusions of the manuscript.

Results

A total of 90 responses generated for 30 standardized clinical questions across the three LLMs were included in the analysis. Inter-rater agreement among the three reviewers for expert-rated accuracy was excellent (ICC=0.90); similarly high consensus was observed across all evaluated quality and reliability instruments (modified DISCERN and EQIP). Descriptive statistics for expert-rated accuracy, modified DISCERN, EQIP, FRES, and FKGL scores are presented in Table 1. Gemini had the highest mean expert-rated accuracy and modified DISCERN scores, whereas Copilot had the highest EQIP and FRES scores and the lowest FKGL score.

Expert-rated accuracy was also examined across the five predefined clinical domains. Gemini showed the highest or joint-highest mean accuracy scores across most domains. The domain-specific accuracy scores for diagnostic evaluation, treatment and initial management, surgical considerations, follow-up and recurrence, and patient counseling are presented in Figure 1.

Overall comparisons using the Friedman test demonstrated statistically significant differences among the three models for all the evaluated metrics (Table 2).

Table 1. Descriptive statistics of model performance across evaluation metrics

Metric	ChatGPT	Gemini	Copilot
Expert-rated accuracy*	4.50±0.57	4.80±0.41	4.43±0.50
Modified DISCERN score*	17.0±1.0	19.0±1.7	17.0±1.5
EQIP score*	85.23±4.43	78.20±4.57	89.60±4.01
FRES*	40.29±3.91	36.20±4.07	42.23±3.89
FKGL*	13.16±1.18	13.43±1.21	12.67±1.04

*: Mean ± standard deviation
EQIP: Ensuring quality information for patients, FRES: Flesch reading ease score, FKGL: Flesch-kincaid grade level

Table 2. Comparison of model performance using the Friedman test

Metric	χ^2 (Friedman)	p-value	Effect size (Kendall W)
Expert-rated accuracy	9.36	0.009	0.16
Modified DISCERN	27.58	<0.001	0.46
EQIP	53.97	<0.001	0.90
FRES	16.62	0.00025	0.29
FKGL	12.0	0.0025	0.23

EQIP: Ensuring quality information for patients, FRES: Flesch reading ease score, FKGL: Flesch-kincaid grade level

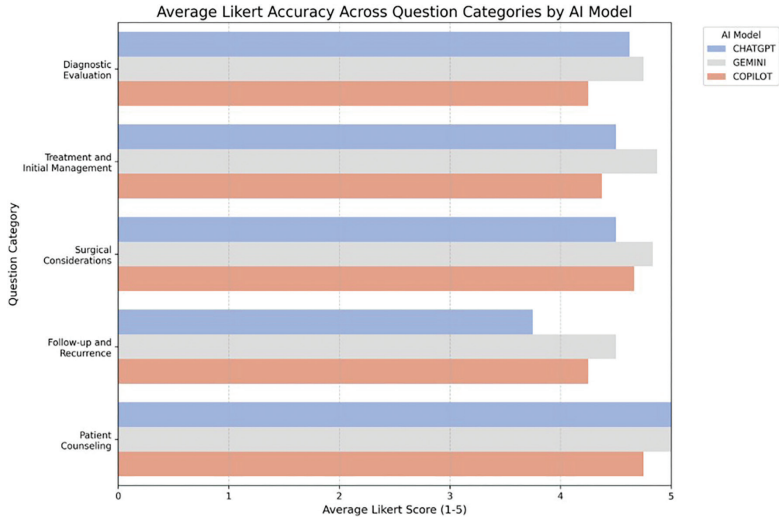


Figure 1. Mean expert-rated accuracy scores of ChatGPT 5.0, Gemini 3.0, and Microsoft Copilot across the five clinical domains evaluated

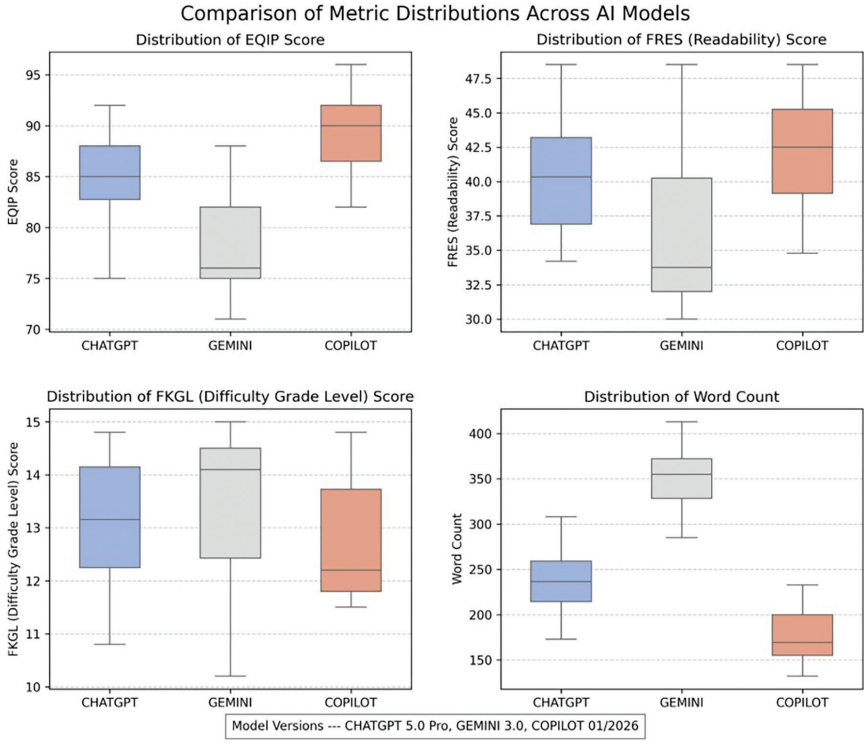


Figure 2. Distribution of modified Word Count, EQIP, FRES, and FKGL scores among ChatGPT 5.0, Gemini 3.0, and Microsoft Copilot
EQIP: Ensuring quality information for patients, FRES: Flesch reading ease score, FKGL: Flesch-kincaid grade level, AI: Artificial intelligence

For the modified DISCERN, a significant between-model difference was observed (χ^2=27.58, p<0.001), with a Kendall's W of 0.46. Significant overall differences were also observed in expert-rated accuracy and the EQIP, FRES, and FKGL scores.

The post-hoc pairwise comparisons are presented in Table 3. Gemini had significantly higher expert-rated accuracy scores than ChatGPT (p=0.012) and Copilot (p=0.0076), whereas no significant difference was observed between ChatGPT and Copilot (p=0.593). The modified DISCERN scores were significantly higher for Gemini than for both ChatGPT and Copilot (both p<0.001), whereas ChatGPT and Copilot did not differ significantly (p=0.924).

Table 3. Pairwise comparisons between models using the Wilcoxon signed-rank test

Metric	ChatGPT vs. Gemini	ChatGPT vs. Copilot	Gemini vs. Copilot
Expert-rated accuracy	0.012	0.593	0.0076
Modified DISCERN	<0.001	0.924	<0.001
EQIP score	<0.001	<0.001	<0.001
FRES	0.00077	0.0041	<0.001
FKGL	0.0095	0.059	0.00096

EQIP: Ensuring quality information for patients, FRES: Flesch reading ease score, FKGL: Flesch-kincaid grade level

For patient-oriented information quality, Copilot had the highest EQIP scores, with significant differences observed among all three model pairs (all p<0.001). Readability also differed among the models. Copilot had the highest FRES and the lowest FKGL values, whereas Gemini had the lowest FRES and the highest FKGL values. Pairwise FRES comparisons were significant for all three model pairs. For FKGL, significant differences were observed between ChatGPT and Gemini and between Gemini and Copilot, whereas the difference between ChatGPT and Copilot was not statistically significant after Bonferroni correction (p=0.059). The distributions of modified DISCERN, EQIP, FRES, and FKGL scores across the three models are shown in Figure 2.

Discussion

The principal finding of the present study was that the three evaluated LLMs demonstrated distinct rather than uniformly superior performance profiles in FUSD-related questions. Gemini achieved the most favorable results for expert-rated accuracy and information reliability, whereas Copilot performed better in patient-oriented information quality and readability. This divergence is clinically relevant because factual or guideline-concordant content does not necessarily translate into information that is easily understood by patients. Accordingly, the evaluation of medical LLMs should extend beyond a single measure of accuracy and incorporate complementary dimensions such as reliability, information quality, and linguistic accessibility.

Our findings are broadly consistent with recent studies that have demonstrated substantial model-dependent variability in AI-generated urological information. Alshak et al. (7) evaluated ChatGPT responses to 125 frequently asked questions across several reconstructive urology domains and reported generally useful information while identifying limitations in information quality and readability. Similarly, a recent review of AI-based patient education in urology concluded that chatbot responses frequently achieve acceptable levels of accuracy but often remain above recommended reading levels (3). More recently, Deger et al. (8) compared multiple AI chatbots for overactive bladder-related patient questions and identified significant differences among models in information quality and readability. These observations support our finding that the relative performance of an LLM depends on the outcome being assessed rather than reflecting a single hierarchy of overall model quality.

A recent guideline-based assessment of contemporary LLMs in preoperative counseling for robot-assisted radical prostatectomy provides an additional comparison. In that study, first-generation responses were independently evaluated by blinded urologists using several complementary measures, including factual accuracy, DISCERN, EQIP, and readability indices. Significant differences were observed across models, and no single quality dimension adequately characterized the overall performance (10). Although the disease context and evaluated models differed from those in the present study, the multidimensional pattern is comparable to our findings and reinforces the importance of evaluating accuracy, information quality, and readability separately when assessing LLM-generated medical information.

This issue may be particularly important in FUSD. Female urethral stricture is an uncommon and heterogeneous condition characterized by non-specific symptoms, variable diagnostic criteria, and several competing therapeutic approaches. Recent reviews have emphasized the continued lack of standardization in its definition, diagnostic pathways, selection of reconstructive techniques, and outcome reporting (2,11). Such heterogeneity creates a demanding context for LLMs because an apparently plausible response may still omit clinically important qualifications regarding diagnostic confirmation, previous interventions, stricture characteristics, or reconstructive options. The higher expert-rated accuracy and modified DISCERN scores observed with Gemini therefore suggest stronger concordance with the reference framework used in this study; however, these findings should not be interpreted as evidence that the model can independently provide individualized clinical decision-making.

Readability represented a separate limitation across the evaluated models. Although Copilot showed the most favorable readability profile, the mean FKGL values for all three models remained approximately at the 13th-grade level, indicating that the generated responses would require relatively advanced reading ability. This observation is consistent with previous urological AI literature, which shows that medically accurate chatbot responses may still be too linguistically complex for routine patient education (3,8). Therefore, optimization of LLM-based patient information should not focus exclusively on factual correctness. Explicit prompting for plain language, shorter sentences, explanation of technical terminology,

and adaptation to the patient's health literacy level may be necessary before such content is used for counseling.

The differences observed between expert-rated accuracy, modified DISCERN, EQIP, and readability also have implications for how these tools might be used in clinical practice. A model that performs well as a professional information resource may not necessarily provide the most suitable patient-facing explanation, and conversely, highly readable responses should not be assumed to have greater guideline concordance. Rather than identifying a universally "best" model, our findings support a task-specific approach to LLM use in which generated information is selected according to its intended purpose and verified by a clinician when applied to a rare and technically nuanced condition, such as FUSD. Given the rapid evolution of LLMs, these results should also be regarded as a comparison of the specific model versions evaluated at the time of the study rather than as a permanent ranking of the platforms.

Study Limitations

This study had several limitations. First, only three LLMs were evaluated at a single time point using versions available in January 2026. Given the rapid evolution of generative AI systems, model updates may affect their performance. Therefore, the present findings should be interpreted as a temporal snapshot rather than a permanent ranking of the platforms. Second, all responses were generated in English using standardized prompts, which may limit the generalizability of the findings to other languages, alternative prompt formulations, and different user-interaction styles. Third, although the responses were independently evaluated by three blinded board-certified urologists and the inter-rater agreement was high, expert-based scoring retained an inherent degree of subjectivity. Fourth, only the first response generated for each question was evaluated, and repeated generations were not performed for the same question. Because LLM outputs may vary across repeated queries, this study did not assess within-model response variability or reproducibility. Fifth, this study evaluated the informational characteristics of AI-generated responses rather than their effects on real-world patient comprehension, clinical decision-making, treatment selection, or patient outcomes. Finally, the analysis focused exclusively on FUSD, and the findings may not be directly generalizable to other areas of reconstructive and functional urology.

Despite these limitations, this study has several strengths. All three LLMs were evaluated using an identical set of standardized questions covering the major clinical domains of FUSD, which allowed direct paired comparisons between models. Responses were independently assessed by blinded board-certified urologists, and inter-rater agreement was formally evaluated. In addition, model performance was examined using complementary measures of expert-rated accuracy, information reliability, patient-oriented information quality, and readability rather than relying on a single outcome measure. Finally, inclusion of the complete question set, standardized prompts, and guideline/reference mapping as supplementary material enhances the transparency and reproducibility of the study.

Conclusion

In this comparative cross-sectional study, the evaluated LLMs demonstrated distinct performance profiles in responding to standardized questions on FUSD. Gemini achieved the highest expert-rated accuracy and modified DISCERN scores, whereas Microsoft Copilot demonstrated the most favorable patient-oriented information quality and readability. These findings indicate that no single model was uniformly superior across all evaluated domains. Therefore, LLM-generated information on FUSD should be interpreted according to its intended use and verified against current clinical guidance before being incorporated into patient education or professional information support.

Ethics

Ethics Committee Approval: This study did not involve human participants, animals, patient data, or biological materials. The study exclusively evaluated responses generated by publicly available artificial intelligence platforms. Therefore, according to institutional and international ethical standards, ethics committee approval was not required.

Informed Consent: Not applicable. The manuscript does not contain any individual person's data, images, or videos; hence, consent for publication was not required.

Authorship Contributions

Surgical and Medical Practices: M.S., S.S., H.L.C., Concept: C.S., Y.C.F., H.L.C., Design: C.S., S.S., Data Collection or Processing: Y.C.F., Analysis or Interpretation: C.S., M.S., Y.C.F., Literature Search: C.S., M.S., H.L.C., Writing: C.S., M.S., S.S.

Conflict of Interest: No conflicts of interest were declared by the authors.

Financial Disclosure: This study received no financial support.

Declaration on the Use of Artificial Intelligence (AI): No artificial intelligence tools were used in the preparation of this manuscript.

Supplementary Table 1 Link: https://d2v96fxpocvxx.cloudfront.net/87bf8cdb-e327-4a57-b2b5-56e4c99c141f/content-images/bea30df0-bdbc-4434-8ceb-508033591400.pdf

References

1. Lumen N, Campos-Juanatey F, Dimitropoulos K et al. European Association of Urology Guidelines on Urethral Stricture Disease. ISBN 978-94-92761-32-5. EAU Guidelines Office, Arnhem, the Netherlands, 2025.
2. Karapanos L, Beck M, Porres D; Trauma and Reconstructive Urology Working Party of the European Association of Urology Young Academic Urologists. Female urethral stricture disease: advances in diagnosis, surgical management, and standardized outcome reporting. Curr Opin Urol. 2026;36:354-62.
3. Paluszek O, Loeb S. Artificial intelligence and patient education. Curr Opin Urol. 2025;35:219-23.
4. Kim ES, Eun SJ, Youn S. The current state of artificial intelligence application in urology. Int Neurourol J. 2023;27:227-33.
5. Gabriel J, Shafik L, Alanbuki A, Larner T. The utility of the ChatGPT artificial intelligence tool for patient education and enquiry in robotic radical prostatectomy. Int Urol Nephrol. 2023;55:2717-32.
6. Ayrancı A, Çağlar U, Yıldız O, Meriç A, Sarılar Ö, Özgör F. Assessing the performance of ChatGPT in addressing questions regarding urethral strictures: a cross-sectional study. J Reconstr Urol. 2024;14:16-22.
7. Alshak MN, Cecelic J, Florissi I, Alam R, Cohen AJ. Assessing ChatGPT responses to frequently asked patient questions in reconstructive urology. Urol Pract. 2025;12:451-8.
8. Deger M, Ates T, Yilmaz IO, et al. Readability and quality assessment of AI-powered chatbot responses to overactive bladder patient questions: a comparative study. BMC Urol. 2026;26:31.
9. Youn S, Park BJ. Artificial intelligence for predicting treatment failure in neurourology: from automated urodynamics to precision management. Int Neurourol J. 2025;29 (Suppl 2):S55-S64.
10. Wang L, Liu H, Yang J, et al. Guideline-based evaluation of large language models for patient counselling before robot-assisted radical prostatectomy: accuracy, safety, and readability. J Robot Surg. 2026;20:888.
11. Turchi B, Lumen N, Verla W, Waterloos M. Female urethral stricture disease: a narrative review on diagnosis, surgical techniques and outcomes. Int J Impot Res. 2026;38:286-95.

Case Report

DOI: 10.4274/haseki.galenos.2026.29494
Med Bull Haseki 2026;64(5):449-453

Atypical Presentation of Primary Central Nervous System Lymphoma in a Young Adult: A Case Report and Current Literature Review

Merve Dur Ince[1], Duygu Turkbey Simsek[2], Ferhat Cuce[3]

[1]*Kahramankazan State Hospital, Clinic of Radiology, Ankara, Türkiye*
[2]*University of Health Sciences Türkiye, Gulhane Training and Research Hospital, Clinic of Radiology, Ankara, Türkiye*
[3]*University of Health Sciences Türkiye, Gulhane Training and Research Hospital, Clinic of Pathology, Ankara, Türkiye*

Abstract

Primary central nervous system lymphoma (PCNSL) is exceptionally rare in the second decade of life and frequently poses a diagnostic challenge when distinguishing PCNSL from high-grade gliomas. We report an 18-year-old male with multifocal brain lesions, in whom the initial mass effect and localization suggested a glial malignancy. However, radiological progression revealed the "vanishing tumor" phenomenon, a hallmark of PCNSL and a critical distinguishing feature from gliomas. Histopathology established a diagnosis of high-grade B-cell lymphoma with an 85% Ki-67 index. This case emphasizes the inclusion of PCNSL in the differential diagnosis of adolescent CNS tumors, as identifying this entity is vital for prioritizing chemotherapy and biopsy over radical surgical resection.

Keywords: Central nervous system, lymphoma, brain

Introduction

Primary central nervous system lymphoma (PCNSL) is a rare and aggressive extranodal non-Hodgkin lymphoma that manifests in the brain parenchyma, leptomeninges, spinal cord, cranial nerves, or intraocular structures without evidence of systemic involvement, accounting for approximately 2-4% of all primary brain tumors (1,2). It constitutes a spectrum of CNS lymphomas occurring in patients with either hereditary or acquired immunodeficiency (3,4). The disease typically presents in individuals aged 60 and older; its occurrence in children and young adults is exceedingly rare and poses significant diagnostic challenges in this age group (4).

Given the clinical rarity and the potential for radiological mimicry among younger patients, we report a case involving an 18-year-old immunocompetent patient diagnosed with PCNSL, focusing on atypical imaging characteristics.

This case report has been prepared in accordance with the CARE (CAse REport) guidelines. Written informed consent was obtained from the patient for the publication of this case report and any accompanying images.

Case Report

An 18-year-old male was admitted to our clinic with a 9-month history of progressive gait disturbance, followed by the onset of dysarthria. The patient's frequent falls, which began at the onset of symptoms, were initially attributed to orthopedic causes. Before admission to our hospital, a lumbar puncture performed at an external center revealed protein levels of 44 mg/dL and glucose levels of 56 mg/dL. Autoimmune limbic encephalitis was suspected, and the patient received intravenous methylprednisolone for 5 days and a cumulative dose of 2 g/kg intravenous immunoglobulin, which was administered as 0.4 g/kg/day over 5 consecutive days (total dose: 150 g). There was no change in the patient's clinical condition during this period. During this clinical course, the patient also experienced generalized tonic-clonic seizures. His medical history was unremarkable except for a known thalassemia

Corresponding Author: Merve Dur Ince, MD, Kahramankazan State Hospital, Clinic of Radiology, Ankara, Türkiye
E-mail: drmervedurince@gmail.com **ORCID:** orcid.org/0009-0008-8162-7222
Received: 06.01.2026 **Accepted:** 17.07.2026 **Publication Date:** 17.11.2026

Cite this article as: Dur Ince M, Turkbey Simsek D, Cuce F. Atypical presentation of primary central nervous system lymphoma in a young adult: a case report and current literature review. Med Bull Haseki. 2026;64(5):449-453

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

trait. The family revealed that his older brother had febrile convulsions and that his parents were first-degree relatives.

Neurological examination: Assessment revealed horizontal gaze-evoked nystagmus. Deep sensation was preserved, but deep tendon reflexes were brisk in the lower extremities. The patient's gait was wide-based and ataxic.

Laboratory and initial imaging workup: Complete blood count and routine biochemical tests were within normal limits. Initial brain magnetic resonance imaging (MRI) demonstrated white matter lesions, predominantly localized in the left frontal lobe. Electroencephalography showed abnormalities characterized by infrequent sharp-wave-slow paroxysms in the right frontotemporoparietal region and similar paroxysms in the left frontotemporal region. Lumbar puncture was performed. Cerebrospinal fluid analysis revealed no cells, but a mild protein elevation (61.7 mg/dL) was noted. Serological screenings for HIV 1-2, hepatitis B (HBsAg), and hepatitis C (anti-HCV) were negative. Complete blood count revealed no lymphopenia, and both serum immunoglobulin levels and CD4+/CD8+ T-lymphocyte ratios were within normal reference ranges. Infectious markers were negative. Oligoclonal bands were positive, showing a Type 4 pattern. Paraneoplastic and autoimmune encephalitis panels were negative. Rheumatological workup showed no abnormalities. Metabolic screening, including very long-chain fatty acids, carnitine-acyl-carnitine analysis, biotinidase activity, and plasma amino acid levels, yielded normal results. Electromyography showed no evidence of neuropathy or myopathy.

Follow-up and progression: A contrast-enhanced follow-up MRI (Figure 1) revealed extensive enhancement in the deep white matter of the left frontal lobe, the left basal ganglia, and the left half of the genu of the corpus callosum.

Within one month, the patient was readmitted because of worsening speech and increasing instability. A repeat brain MRI (Figure 2) demonstrated new contrast-enhancing lesions with diffusion restriction in the right hemisphere, while the size and enhancement of the previously noted left-sided lesions had decreased, likely reflecting a transient response to prior corticosteroid therapy. Consequently, digital subtraction angiography was performed to evaluate for vasculitis, and whole-exome sequencing (WES) was initiated. Digital subtraction angiography and WES results were unremarkable. A biopsy of the right frontal lobe was planned. The clinical materials and diagnostic data for this case report were obtained in February 2025.

Histopathological findings: Examination of sections from the right frontal lobe biopsy revealed an atypical lymphoid cell population infiltrating the brain parenchyma, characterized by blastic chromatin, prominent apoptosis, and increased mitotic activity (Figure 3). Immunohistochemical analysis showed that the tumor cells were diffusely positive for CD20, PAX-5, Bcl-6, CD10, and c-MYC, while they were negative for Bcl-2, CD30, and ALK (Figure 4). Within and adjacent to the tumor, a population of CD3-positive reactive T-lymphoid cells was observed, concentrating in perivascular areas and extending into the parenchyma (Figure 5).

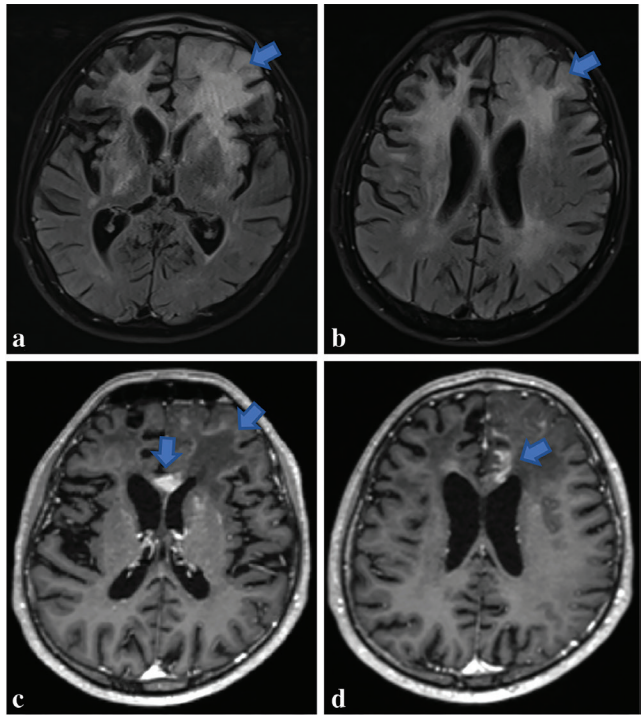


Figure 1. (a,b) Non-contrast brain MRI, January 2025: Axial FLAIR sequences demonstrate hyperintense lesions in the periventricular white matter of both frontal lobes and within the basal ganglia. A voluminous lesion involving the cortex and white matter is noted in the left frontal lobe, causing effacement of the adjacent cortical sulci (blue arrow). **(c,d)** Post-contrast T1-weighted images reveal peripheral linear enhancement in the anterior segment of the left frontal lesion, along with nodular and punctate enhancement at the level of the corpus callosum and the left caudate nucleus (blue arrows)
MRI: Magnetic resonance imaging, FLAIR: Fluid-attenuated inversion recovery

Discussion

Primary central nervous system lymphoma is an aggressive malignancy accounting for approximately 1-5% of all primary brain tumors and nearly 1% of all non-Hodgkin lymphomas (1,2). Although it predominantly affects immunocompromised individuals, its occurrence in immunocompetent young adults is rare and represents an atypical presentation that poses a significant clinical and

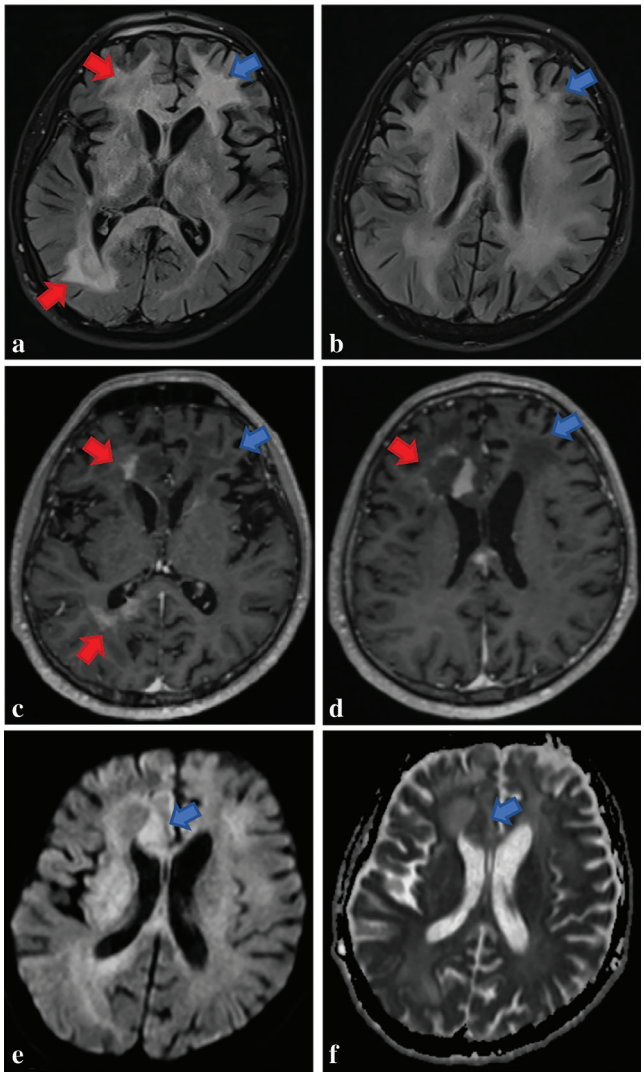


Figure 2. (a,b) Axial FLAIR sequence from February 2025 demonstrates regression in the size of the previously identified lesion in the anterior segment of the left frontal lobe (blue arrow). Concurrently, new lesions with significant mass effect and sulcal effacement have developed in the right frontal lobe adjacent to the corpus callosum (involving the cortico-medullary junction) and within the right periatrial white matter (red arrows). **(c,d)** Post-contrast images reveal ill-defined (faint) peripheral enhancement in the right frontal lesion and heterogeneous enhancement in the right periatrial white matter lesion (red arrows). The regressing lesion in the left frontal lobe shows no contrast enhancement (blue arrow). **(e,f)** Diffusion-weighted imaging shows hyperintensity along the medial aspect of the right frontal lesion, with corresponding mild hypointensity on the apparent diffusion coefficient map, indicating restricted diffusion. **(a,d)** In both January and February MRI studies, the cortical sulci were prominent, reflecting a decrease in brain parenchymal volume (atrophy) disproportionate to the patient's age

MRI: Magnetic resonance imaging, FLAIR: Fluid-attenuated inversion recovery

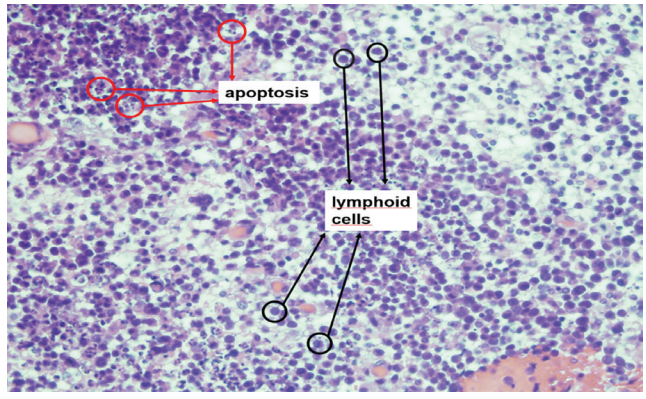


Figure 3. Histopathological examination demonstrates a monotonous population of medium-to-large lymphoid cells infiltrating the brain parenchyma. The cells exhibit blastic chromatin, scant cytoplasm, prominent apoptosis, and increased mitotic activity (H&E staining, ×400)

H&E: Hematoxylin and eosin

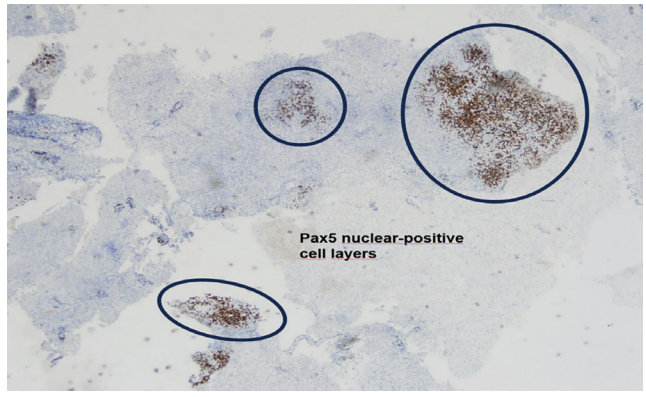


Figure 4. Photomicrograph showing diffuse and strong nuclear immunoreactivity for PAX5 within the neoplastic cells previously identified in Figure 3 (x40 magnification)

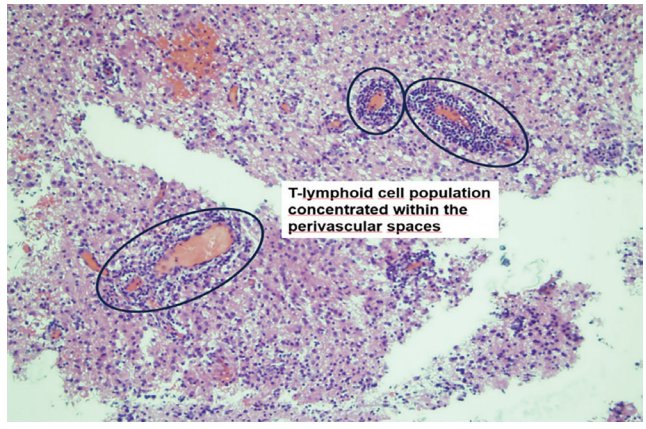


Figure 5. Photomicrograph demonstrates a T-lymphoid cell population concentrated within the perivascular spaces and extending into the brain parenchyma, observed both within the tumor and in the adjacent tissues (H&E, x200 magnification)

H&E: Hematoxylin and eosin

radiological diagnostic challenge (3,4).

The differential diagnosis for PCNSL is notoriously broad, encompassing secondary lymphoma, glioblastoma, metastatic disease, viral or autoimmune encephalitis, and tumefactive demyelinating lesions (5,6). In the CNS, PCNSL primarily manifests as infiltrative parenchymal lesions. While multifocality and necrosis are more common among immunocompromised patients, immunocompetent patients often present with solitary, non-necrotic masses. Predilection sites include the periventricular white matter, basal ganglia, and corpus callosum (7,8). On MRI, PCNSL typically appears iso- to hypointense on both T1- and T2-weighted sequences, a key reflection of its high nuclear-to-cytoplasmic ratio and compact cellularity (7,8). A hallmark of the disease is marked diffusion restriction on apparent diffusion coefficient (ADC) maps, which serves as a critical differentiator from high-grade gliomas; PCNSL typically exhibits higher cellular density and lower ADC values than glial tumors (5,6). Following contrast administration, intense and homogeneous enhancement is the classic presentation, reflecting blood-brain barrier disruption (6).

The radiological findings in this 18-year-old patient presented a significant diagnostic challenge, with an initial differential diagnosis encompassing viral or autoimmune encephalitis, tumefactive demyelinating lesions, and high-grade gliomas. The initial MRI, performed at an external facility in December 2024, revealed multifocal FLAIR hyperintensities involving the periventricular white matter and basal ganglia, suggesting an inflammatory or infectious process. However, the subsequent spatiotemporal evolution of the lesions provided the first critical clue for PCNSL. The regression of the left frontal lesion within one month, a classic manifestation of the "vanishing" or "ghost tumor" phenomenon, is highly characteristic of PCNSL's sensitivity to corticosteroids or its naturally fluctuating course, a feature rarely seen in gliomas or metastatic disease (2,9). This transient response significantly hinders definitive diagnosis and delays essential treatment. Therefore, in cases where clinico-radiological discordance exists, early stereotactic biopsy is mandatory to circumvent diagnostic delays.

Furthermore, the emergence of new lesions in the right frontal and periatrial regions (February 2025) with significant mass effect and restricted diffusion (hyperintensity on diffusion-weighted imaging and mild hypointensity on ADC) served as a key radiologic differentiator. Restricted diffusion is a hallmark of the high cellularity (hypercellularity) typical of lymphoma, whereas autoimmune encephalitis typically lacks such focal restricted diffusion and rarely presents with the heterogeneous/peripheral enhancement patterns observed in this case (6,10). Additionally, while tumefactive demyelinating lesions may show peripheral enhancement, they typically exhibit an "open-ring" pattern and lack the dense, solid hypercellularity seen on ADC maps (6,7). Finally, the presence of brain atrophy disproportionate to the patient's age added a layer of complexity, necessitating the exclusion of underlying systemic immunodeficiencies, which were subsequently ruled out, thereby confirming the atypicality of this PCNSL presentation in an immunocompetent young adult (1,4).

To contextualize our findings, Table 1 summarizes recent literature on atypical PCNSL in young or immunocompetent patients. Unlike classic homogeneous masses, these cases increasingly exhibit atypical features such as the 'vanishing tumor' phenomenon or encephalitis-like presentation. This underscores the critical need to consider PCNSL in the differential diagnosis of atypical CNS lesions in young adults, regardless of immune status.

Histopathologically, while the majority of PCNSL cases are diffuse large B-cell lymphomas, our case presented a unique diagnostic conundrum (1,5). Initial immunohistochemical markers (specifically CD10 and Bcl-6 positivity and widespread c-MYC positivity) strongly raised suspicion for Burkitt lymphoma. However, a Ki-67 index of 85% and a prominent T-lymphoid population intermingled with the neoplastic cells were more consistent with high-grade large B-cell lymphoma. This dense, reactive cellular infiltrate, combined with the patient's protracted clinical history, suggests a malignancy arising within an autoinflammatory background. Such a complex reactive environment likely acted as a histopathological masquerade, contributing to the initial diagnostic confusion and the "encephalitis-like" presentation that defined the early clinical course (9). Ultimately, the absence of EBV-LMP1 and ALK expression further refined the diagnosis, confirming an aggressive PCNSL subtype in an immunocompetent host (1,3).

Table 1. Summary of recent literature highlighting atypical presentations and diagnostic challenges of Primary Central Nervous System Lymphoma (PCNSL)

Author, year (ref)	Study design	Main focus/atypical feature
Van Der Meulen et al. (2), 2026	Review	Current standards and developments
Ayalew et al. (4), 2024	Case report	Young immunocompetent patient
Ribas et al. (6), 2024	Review	Imaging features and differential diagnosis
Zawar et al. (10), 2021	Case report	Mimicking herpes encephalitis

Conclusion

Primary central nervous system lymphoma is a rare and aggressive malignancy that can mimic glioblastoma, demyelinating lesions, encephalitis, and other neurological disorders, even in patients during the second decade of life. In patients presenting with atypical CNS lesions and negative viral, autoimmune, or metabolic workups, a fluctuating radiological pattern (characterized by the paradoxical regression of initial lesions and the emergence of new ones) should strongly heighten clinical suspicion for lymphoma. This case highlights the critical necessity of avoiding corticosteroid administration before a tissue biopsy. High-potency corticosteroids can induce rapid cytoreduction in PCNSL, leading to the "vanishing tumor" phenomenon, which may preclude a definitive histological diagnosis and delay life-saving treatment.

This case adds to the limited literature on PCNSL in young immunocompetent patients by emphasizing that a fluctuating radiological course ("vanishing tumor" phenomenon) may represent an important diagnostic clue. Early recognition of this pattern may shorten diagnostic delay, prevent inappropriate surgical management, and facilitate prompt initiation of lymphoma-directed therapy.

Ethics

Informed Consent: Written informed consent was obtained from the patient for the publication of this case report and any accompanying images.

Footnotes

Authorship Contributions

Concept: M.D.I., F.C., Design: M.D.I., F.C., Data Collection or Processing: M.D.I., D.T.S., F.C., Analysis or Interpretation: M.D.I., D.T.S., F.C., Literature Search: M.D.I., D.T.S., Writing: M.D.I., D.T.S., F.C.

Conflict of Interest: No conflicts of interest were declared by the authors.

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

References

1. Ferreri AJM, Calimeri T, Cwynarski K, et al. Primary central nervous system lymphoma. Nat Rev Dis Primers. 2023;9:29.
2. Van Der Meulen M, Vermaat JSP, Galldiks N, et al. Primary central nervous system lymphoma-current standards and recent developments. Neurooncol Pract. 2026;13:640-8.
3. Deckert M, Engert A, Brück W, et al. Modern concepts in the biology, diagnosis, differential diagnosis and treatment of primary central nervous system lymphoma. Leukemia. 2011;25:1797-807.
4. Ayalew ZS, Gebregiorgis M, Azibte GT, Hamza AK, Abdo IS, Molla BA. Primary central nervous system lymphoma: a diagnostic challenge in a young immunocompetent patient with limited resources. Radiol Case Rep. 2024;19:4644-9.
5. Han Y, Wang ZJ, Li WH, et al. Differentiation between primary central nervous system lymphoma and atypical glioblastoma based on MRI morphological feature and signal intensity ratio: a retrospective multicenter study. Front Oncol. 2022;12:811197.
6. Ribas GA, de Mori LH, Freddi TAL, Oliveira LDS, de Souza SR, Corrêa DG. Primary central nervous system lymphoma: imaging features and differential diagnosis. Neuroradiol J. 2024;37:705-22.
7. Haldorsen IS, Espeland A, Larsson EM. Central nervous system lymphoma: characteristic findings on traditional and advanced imaging. AJNR Am J Neuroradiol. 2011;32:984-92.
8. Slone HW, Blake JJ, Shah R, Guttikonda S, Bourekas EC. CT and MRI findings of intracranial lymphoma. AJR Am J Roentgenol. 2005;184:1679-85.
9. Feldheim J, Darkwah Oppong M, et al. Histopathological diagnosis of primary central nervous system lymphoma after therapy with corticosteroids or anticoagulants. Cancers (Basel). 2024;16:1157.
10. Zawar I, Cho SM, Prayson R, Levin K. Primary central nervous system lymphoma mimicking herpes encephalitis. Neurooncol Pract. 2021;8:620-2.

Case Report

DOI: 10.4274/haseki.galenos.2026.08370
Med Bull Haseki 2026;64(5):454-457

Retrograde Bypass Due to Chronic Mesenteric Ischemia after Left Ventricular Aneurysm in a Patient with Woven Coronary Artery Anomaly: A Case Report and Current Literature Review

Osman Fehmi Beyazal[1], Bilge Ecemis Yilmaz[2], Korhan Erkanli[2]

[1]*University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Cardiovascular Surgery, Istanbul, Türkiye*
[2]*Medipol Mega University Hospital, Department of Cardiovascular Surgery, Istanbul, Türkiye*

Abstract

A 64-year-old man presented with a 3-month history of postprandial abdominal pain and weight loss. Abdominal computed tomographic angiography (CTA) revealed a long-segment total occlusion of the superior mesenteric artery (SMA). Chest CTA incidentally detected a 3-cm left ventricular (LV) aneurysm. Coronary angiography demonstrated 80% stenosis of the left anterior descending (LAD) artery and a woven coronary artery anomaly involving the distal left circumflex artery. Following median sternotomy, the 6 × 3-cm LV aneurysm was resected, and the intraluminal thrombus was removed. Coronary artery bypass grafting to the LAD artery was then performed using the left internal mammary artery. Fourteen days later, a second operation was performed, during which a saphenous vein graft was used to create a retrograde bypass between the right common iliac artery and the SMA. This case highlights the importance of comprehensive cardiac evaluation in patients with chronic mesenteric ischemia and demonstrates that staged retrograde SMA bypass can be performed safely in selected patients.

Keywords: Mesenteric Ischemia, aneurysm, coronary vessel anomalies

Introduction

Mesenteric artery stenosis (MAS) results in inadequate blood flow to the gastrointestinal tract, leading to chronic mesenteric ischemia (CMI), which significantly impairs patients' quality of life (1). The rich and complex collateral system within the intestines protects against potential ischemia (2). This is particularly important in patients with occlusion of more than one mesenteric artery. Chronic mesenteric ischemia may occur due to factors such as atrial fibrillation (AF), which can lead to mesenteric artery embolism; atherosclerosis, which can cause mesenteric artery thrombosis; and, rarely, cardiac aneurysms with embolization of formed thrombi (1). We present the case of a patient with a woven coronary artery (WCA) anomaly in whom CMI developed after a left ventricular (LV) aneurysm.

Case Report

A 64-year-old male patient, weighing 48 kg, presented with postprandial abdominal pain and weight loss over the past 3 months. Apart from a history of polycythemia vera and benign prostatic hyperplasia, no other medical conditions were reported. He had undergone surgery for the latter. Laboratory examinations did not reveal any significant abnormalities. His heart rate was within normal range and he was in sinus rhythm. Abdominal computed tomographic angiography (CTA) revealed a long-segment total occlusion of the proximal superior mesenteric artery

Corresponding Author: Osman Fehmi Beyazal, MD, University of Health Sciences Türkiye, Basaksehir Cam and Sakura City Hospital, Clinic of Cardiovascular Surgery, Istanbul, Türkiye
E-mail: osmanfehmibeyazal@gmail.com **ORCID:** orcid.org/0000-0001-6211-0676
Received: 20.09.2025 **Accepted:** 26.02.2026 **Publication Date:** 17.11.2026

Cite this article as: Beyazal OF, Yilmaz BE, Erkanli K. Retrograde bypass due to chronic mesenteric ischemia after left ventricular aneurysm in a patient with woven coronary artery anomaly: a case report and current literature review. Med Bull Haseki. 2026;64(5):454-457

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

(SMA), with collateral circulation in the distal segment. A 90% short-segment stenosis was present in the celiac trunk orifice, and a 70% stenosis was present in the right renal artery (Figure 1). On thoracic CTA, an aneurysm measuring 3-cm in diameter was incidentally detected on the inferolateral wall at the base of the LV. This aneurysm was also evaluated with cardiac magnetic resonance imaging, and findings consistent with a chronic infarct in the LV lateral segment were observed. Transthoracic echocardiography revealed an LV aneurysm and a thrombus within it. Coronary angiography (CAG) revealed 80% stenosis in the left anterior descending (LAD) artery, 40% stenosis in the right coronary artery, distal stenosis in the left circumflex artery, and a woven coronary artery anomaly involving the distal left circumflex artery. (Figure 1). Despite lacking cardiac symptoms, the patient underwent multidisciplinary evaluation by cardiology, anesthesiology, interventional radiology, general surgery, and cardiovascular surgery. The team decided to prioritize a cardiac operation, followed by an open operation for the SMA. Additionally, a decision was made to place a stent in the celiac artery (CA). Due to the absence of renal dysfunction and hypertension, renal artery intervention was not considered.

In June 2024, a 6 x 3 cm aneurysm was surgically opened and resected, and the intraluminal thrombus was removed (Figure 2). Following this, a bypass procedure was performed on the LAD artery using the left internal mammary artery. The cross-clamp time was 20 minutes, and the cardiopulmonary bypass time was 51 minutes. Only 1 unit of fresh frozen plasma was used; no other blood products were required. The patient was discharged from the intensive care unit on the first postoperative day and had no complications during the follow-up period. The second operation, planned 14 days after the first, was successfully performed. Following a median laparotomy, the mesenteric artery and its branches were located using intraoperative ultrasonography and then surgically marked. A bypass was then created between the right common iliac artery (RCIA) and the SMA using a saphenous vein graft harvested from the right leg in a C-loop configuration (Figure 3). The patient was discharged from the intensive

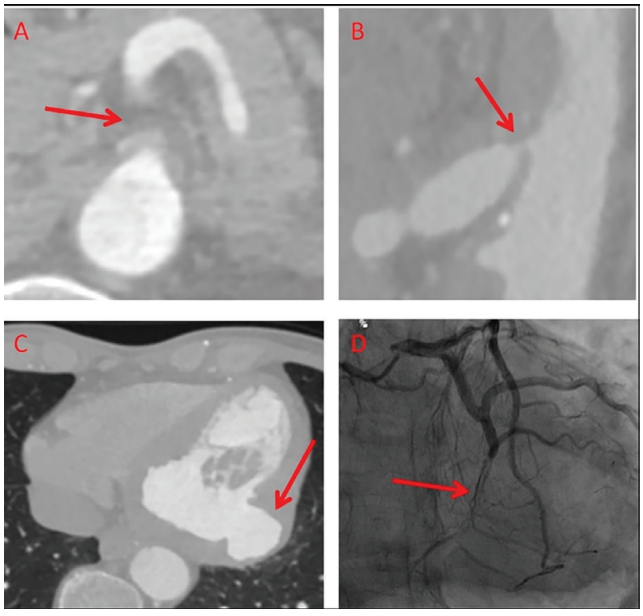


Figure 1. A) Long-segment total occlusion in the proximal part of the SMA, **B)** short-segment stenosis in the celiac trunk orifice, **C)** aneurysm on the inferolateral wall at the base of the LV, **D)** woven coronary artery anomaly involving the distal left circumflex artery
SMA: Superior mesenteric artery

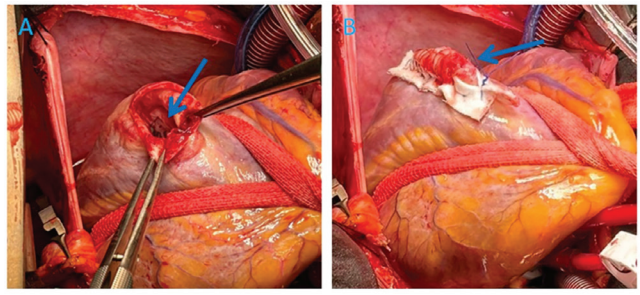


Figure 2. A) Aneurysm that opens linearly and contains thrombus, **B)** intraoperative view after aneurysm repair

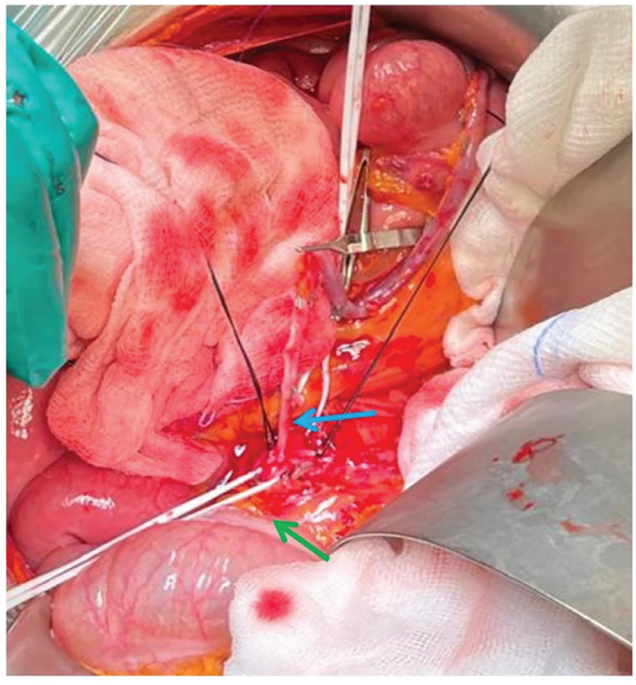


Figure 3. Saphenous vein graft bypassed between the right common iliac artery and superior mesenteric artery. The blue arrow shows the saphenous vein graft. The green arrow shows the superior mesenteric artery where the anastomosis was performed

care unit the day after surgery. Upon complete resolution of the patient's symptoms following the SMA bypass, a multidisciplinary evaluation was conducted to address the CA lesion. It was decided that the patient would undergo regular medical follow-up. After 8 days, the patient was discharged with a prescription for acetylsalicylic acid (100 mg) and experienced no complications. No adverse events occurred during postoperative follow-up. Patient consent was obtained for this case report.

Discussion

Chronic mesenteric ischemia is a pathology characterized by postprandial abdominal pain lasting about 1 hour and weight loss caused by nausea, vomiting, diarrhea, and fear of food (3). The most common cause of mesenteric ischemia is atherosclerosis of the mesenteric arteries. However, it can also occur due to AF and, rarely, due to thrombus formation associated with LV aneurysms, as in this case report (1). Ischemic symptoms often occur when two of the three main mesenteric arteries are diseased (2). Symptomatic patients should be treated. There is no consensus on treatment options, as various methods have been described.

When deciding on interventions for symptomatic patients, it is important to evaluate patients using a multidisciplinary approach. Both endovascular and open surgical options should be discussed. Currently, endovascular treatment, which involves angioplasty and stenting, has shown high success rates and low mortality rates. It is often considered the first-line treatment (3). However, for patients with long segments, consecutive lesions, or complex anatomy who are not suitable for endovascular treatment, open bypass surgery remains a more appropriate treatment option. We present a case of WCA in a patient who developed CMI after an LV aneurysm. To the best of our knowledge, this is the first case report of a patient undergoing bypass surgery for CMI after an LV aneurysm who additionally had a WCA anomaly.

For patients planning to undergo bypass surgery, the target vessels must first be identified. The primary target for revascularization is the SMA. Revascularization of the CA or inferior mesenteric artery may also be considered in patients who are not suitable for SMA revascularization. SMA revascularization alone has been associated with acceptable mortality rates and survival outcomes. Additionally, SMA bypass alone has been shown to be as successful as multivessel bypass in treating bowel ischemia (4). It was reported that a permanent improvement in symptoms was achieved in 73% of patients with CMI due to a single MAS who underwent CA or SMA revascularization (5). In our case, we planned surgery because the lesion in the SMA involved a long segment and was completely occluded, making it unsuitable for endovascular treatment. We decided to place a stent in the CA because the lesion was short-segment, making it more suitable for endovascular treatment. Since the primary target vessel for bowel ischemia was the SMA, we performed an SMA bypass first. To minimize surgical risk, we initially planned to stent the CA after the operation. Although the patient's chronic abdominal pain had completely resolved after the operation, we decided to continue medical treatment for the time being.

Various bypass methods have been described in the literature, with antegrade and retrograde approaches commonly used. In an antegrade bypass, the proximal portion of the graft is connected to the distal thoracic aorta or to the proximal abdominal aorta. In retrograde bypass, the distal abdominal aorta, a previous aortic graft, or the iliac arteries are often employed (6). Although the absence of aortic clamping during retrograde bypass is an advantage, the procedure is usually performed in a C-loop configuration; therefore, avoiding angulation and kinking of the graft is important. We used the retrograde bypass technique with a C-loop configuration to minimize the risks associated with aortic clamping. Despite only an SMA bypass being performed, all of the patient's symptoms were fully resolved, and no further interventions were necessary.

One of the most important aspects of studying this disease is investigating its etiology. While CMI is typically caused by atherosclerosis, it can also result from an LV aneurysm, as seen in this case. Despite reporting no cardiac symptoms, the patient was found to have an LV aneurysm on routine preoperative screening and a lesion in the LAD on subsequent CAG. As a result, the patient required cardiac surgery prior to the SMA bypass. Screening patients with CMI for potential cardiac issues could be beneficial. This case report, detailing the first SMA bypass performed after an LV aneurysm, will contribute significantly to the existing literature.

Another important aspect of this case is the presence of a rare anomaly, WCA. Woven coronary artery is a sporadic congenital anomaly in which the coronary artery divides into thin channels and subsequently re-enters the distal lumen. This condition is typically considered benign and often asymptomatic (7). We did not plan any intervention because the patient was asymptomatic.

Conclusion

Although endovascular methods are currently the preferred treatment for CMI, open surgical options should be considered for symptomatic patients for whom endovascular treatment is not suitable. Retrograde bypass can be successfully performed through the RCIA. CMI rarely develops as a result of a LV aneurysm; therefore,

patients presenting with CMI should undergo a detailed cardiac evaluation.

Ethics

Informed Consent: Patient consent was obtained for this case report.

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.E.Y., K.E., Concept: O.F.B., Design: O.F.B., B.E.Y., K.E., Data Collection or Processing: O.F.B., B.E.Y., K.E., Analysis or Interpretation: O.F.B., B.E.Y., K.E., Literature Search: O.F.B., Writing: O.F.B.

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

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

References

1. Zafarghandi MR, Samimiat A, Nikraftar P, Sadeghi A, Pourazari P, Shokri A. Aortic reimplantation of the superior mesenteric artery (SMA) for SMA stenosis in a previously stented patient: new technique and a case report. Int J Surg Case Rep. 2024;115:109170.
2. Geng B, Menth M, Widmer L, Egger B, Psathas E. Retrograde mesenteric perfusion from the deep femoral artery in a patient with a recurrent anastomotic aneurysm in the groin: a case report. Front Surg. 2023;10:1208871.
3. Jun HJ. Isolated bypass to the superior mesenteric artery for chronic mesenteric ischemia. Korean J Thorac Cardiovasc Surg. 2013;46:146-9.
4. Gentile AT, Moneta GL, Taylor LM Jr, Park TC, McConnell DB, Porter JM. Isolated bypass to the superior mesenteric artery for intestinal ischemia. Arch Surg. 1994;129:926-31; discussion 931-2.
5. van Dijk LJD, Moons LMG, van Noord D, et al. Persistent symptom relief after revascularization in patients with single-artery chronic mesenteric ischemia. J Vasc Surg. 2018;68:779-85.
6. Björck M, Koelemay M, Acosta S, et al. Editor's choice - management of the diseases of mesenteric arteries and veins: clinical practice Guidelines of the European Society of Vascular Surgery (ESVS). Eur J Vasc Endovasc Surg. 2017;53:460-510.
7. Ayhan S, Ozturk S, Tekelioglu UY, Ocak T. Woven coronary artery anomaly associated with acute coronary syndrome. Int J Angiol. 2013;22:55-8.

Letter to the Editor

DOI: 10.4274/haseki.galenos.2026.86547
Med Bull Haseki 2026;64(5):458-459

Distinguishing Sarcopenia From Parkinsonian Motor Dysfunction: Limitations of SARC-F Screening

Okan Sokmen

University of Health Sciences Türkiye, Ataturk Sanatorium Training and Research Hospital, Clinic of Neurology, Ankara, Türkiye

To the editor,

I read with interest the article entitled "Association between sarcopenia risk, disease severity, and functional mobility in Parkinson's disease: a cross-sectional study" (1). The study highlights a clinically significant relationship among sarcopenia risk, disease severity, and functional mobility in Parkinson's disease. However, the interpretation of these findings requires careful consideration from a neurological perspective. The strength, assistance with walking, rising from a chair, climbing stairs, and falls (SARC-F) questionnaire is not a direct measure of muscle mass or muscle quality; rather, it is a screening tool reflecting functional limitations such as gait disturbances, difficulty getting up from a chair, and balance problems (2). These domains substantially overlap with the core motor features of Parkinson's disease, including bradykinesia, rigidity, and postural instability (3). Therefore, high SARC-F scores in this patient group may primarily represent underlying Parkinsonian motor severity rather than true sarcopenia (4).

Current diagnostic approaches emphasize that sarcopenia should not be defined solely on the basis of functional limitations (2). According to contemporary consensus criteria, the diagnosis of sarcopenia requires the combined assessment of muscle strength, muscle quantity or quality, and physical performance (2). In this context, sarcopenia assessment should move beyond questionnaire-based screening to include objective measures such as handgrip strength, chair-stand performance, and muscle mass evaluation using dual-energy X-ray absorptiometry or bioelectrical impedance analysis (2,5). Emerging tools such as muscle ultrasonography may offer additional value in distinguishing true muscle loss from Parkinson-related motor dysfunction (5).

In Parkinson's disease, this multidimensional approach is particularly important to distinguish disease-related functional impairment from true muscle loss (4,6). Recent studies also support this observation. A study showed that the SARC-F questionnaire has limited diagnostic performance for detecting sarcopenia in Parkinson's disease, has low sensitivity, and its performance is similar to disease staging scales (4). Similarly, another study reported that Parkinson's patients showed greater functional decline and higher screening scores. Nevertheless, the sarcopenia prevalence defined according to objective criteria did not differ significantly from control groups (6). In addition, emerging evidence indicates that muscle assessment methods, including ultrasonography, may provide more reliable evaluation of muscle status in this population (5).

In this context, the strong correlations reported in the study (1) may not indicate a direct relationship with sarcopenia. Instead, they may reflect overlapping measurement structures related to Parkinsonian motor burden (4,6). This distinction is clinically crucial because sarcopenia is a potentially modifiable condition through nutritional and resistance-based interventions (7), whereas Parkinsonian motor disorders follow a neurodegenerative process and require different therapeutic strategies (6). From a neurological perspective, this differentiation is not only methodological but also clinically relevant. In Parkinson's disease, a significant portion of functional limitations stems from impaired motor control due to dopaminergic dysfunction in basal ganglia circuits (3), while in sarcopenia, it is primarily associated with structural

Corresponding Author: Okan Sokmen, MD, University of Health Sciences Türkiye, Ataturk Sanatorium Training and Research Hospital, Clinic of Neurology, Ankara, Türkiye
E-mail: drokansokmen@gmail.com **ORCID:** orcid.org/0000-0001-8378-1233
Received: 03.04.2026 **Accepted:** 20.04.2026 **Publication Date:** 17.11.2026

Cite this article as: Sokmen O. Distinguishing sarcopenia from parkinsonian motor dysfunction: limitations of SARC-F screening. Med Bull Haseki. 2026;64(5):458-459

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

and metabolic changes in peripheral muscle tissue (7). These two processes can produce similar functional-level results, which can cause diagnostic overlap, especially in assessments based solely on functional scales. Therefore, integrating direct measurements of muscle strength and muscle mass with neurological examination findings, disease stage, and motor phenotype may provide a more accurate clinical approach to sarcopenia assessment in Parkinson's disease (6,7).

Therefore, the key issue may not be whether sarcopenia risk correlates with mobility, but whether current assessment approaches can meaningfully differentiate sarcopenia from Parkinson-related motor dysfunction. In Parkinson's disease, sole reliance on functional screening tools may be insufficient, and a more comprehensive approach supported by objective muscle measurements may be required.

Footnotes

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

References

1. Asik HK, Zure M, Oguzhanasilturk DU, Gungor ZK. Association between sarcopenia risk, disease severity, and functional mobility in Parkinson's disease: a cross-sectional study. Med Bull Haseki. 2026;64:147-54.
2. Cruz-Jentoft AJ, Bahat G, Bauer J, et al. Sarcopenia: revised European consensus on definition and diagnosis. Age Ageing. 2019;48:16-31.
3. Tanner CM, Ostrem JL. Parkinson's disease. N Engl J Med. 2024;391:442-52.
4. Choi S, Kim R, Lee JY, Jeon B. Clinical usefulness of the SARC-F questionnaire to assess sarcopenia in patients with Parkinson's disease. Parkinsonism Relat Disord. 2024;125:107039.
5. Lim LS, Lu CH, Lai YR, Kan NN, Liao CC, Lee SY. Muscle ultrasonography as a diagnostic tool for assessing sarcopenia in Parkinson's disease. J Mov Disord. 2025;18:347-54.
6. Gomes de Luna JR, Lima DP, Gomes VC, et al. Sarcopenia, body composition, physical performance, and clinical features in Parkinson's disease: a cross-sectional comparative study. Front Endocrinol. 2026;17:1793254.
7. Cruz-Jentoft AJ, Sayer AA. Sarcopenia. Lancet. 2019;393:2636-46.

Letter to the Editor

DOI: 10.4274/haseki.galenos.2026.92905
Med Bull Haseki 2026;64(5):460-461

The Importance of Structured Neuropathic Symptom Assessment in Kinesio Taping Studies on Carpal Tunnel Syndrome

Taner Dandinoglu, Erkan Kaya

University of Health Sciences Türkiye, Bursa City Hospital, Clinic of Physical Medicine and Rehabilitation, Bursa, Türkiye

To the Editor,

Kinesio taping is widely used as a supportive intervention across various clinical conditions, particularly musculoskeletal disorders. For this reason, evaluating a commonly applied technique for one of the most prevalent peripheral nerve entrapments is both encouraging and clinically relevant. We found the study engaging for clinicians and promising in its results, and we congratulate the authors on their valuable contribution (1).

Carpal tunnel syndrome (CTS) is, by its very nature, a condition in which neuropathic symptoms form a substantial part of the clinical picture. Many patients experience difficulty with daily activities, specifically owing to these neuropathic features. Importantly, these symptoms are not limited to daytime functioning; they frequently become more pronounced at night, when patients seek rest after the day's fatigue, and this exacerbation can significantly disrupt sleep quality (2,3).

Systematic evaluation and follow-up of neuropathic symptoms using validated instruments such as the DN4 or LANSS questionnaires, applied at baseline and predefined post-intervention follow-up time points, may provide clinicians with a clearer and more objective basis for treatment decisions. For these reasons, we believe that, particularly in interventional studies, evaluating neuropathic symptoms with validated, reliable instruments designed for this purpose may not only enhance the objectivity of the study but also increase its clinical relevance and strengthen the credibility of its findings (4,5).

We intend not to diminish the value of the authors' work, but rather to highlight an aspect that we believe could enrich future research in this field. Incorporating structured neuropathic assessments into Kinesio taping studies may help clarify which sensory domains are most influenced by the intervention and may provide clinicians with a more comprehensive framework for interpreting treatment outcomes. We once again commend the authors for their contribution and hope that this perspective will encourage continued refinement of assessment strategies in conservative CTS management.

Footnotes

Authorship Contributions

Concept: T.D., E.K., Design: T.D., E.K., Data Collection or Processing: T.D., E.K., Analysis or Interpretation: T.D., E.K., Literature Search: T.D., E.K., Writing: T.D., E.K.

Conflict of Interest: No conflicts of interest were declared by the authors.

Financial Disclosure: This study received no financial support.

Corresponding Author: Taner Dandinoglu, MD, University of Health Sciences Türkiye, Bursa City Hospital, Clinic of Physical Medicine and Rehabilitation, Bursa, Türkiye
E-mail: dandinoglu@gmail.com **ORCID:** orcid.org/0000-0002-6010-7315
Received: 22.11.2025 **Accepted:** 26.02.2026 **Publication Date:** 17.11.2026

Cite this article as: Dandinoglu T, Kaya E. The importance of structured neuropathic symptom assessment in kinesio taping studies on carpal tunnel syndrome. Med Bull Haseki. 2026;64(5):460-461

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

References

1. Baykal Sahin H, Kara F, Cakir E, Saglam Akkaya G. Effects of Kinesio taping in carpal tunnel syndrome treatment: a randomized controlled trial. Med Bull Haseki. 2025;63:171-7.
2. Gautam J, Seshagiri DV, Haldar P, et al. Sleep profile and its clinical correlates in patients with carpal tunnel syndrome. Neurology Asia. 2025;30:771-80.
3. Đilvesi Đ, Jelača B, Knežević A, Živanović Ž, Pantelić V, Golubović J. Evolution of carpal tunnel syndrome treatment: a narrative review. NeuroSci. 2026;7:10.
4. Padua L, Coraci D, Erra C, et al. Carpal tunnel syndrome: clinical features, diagnosis, and management. Lancet Neurol. 2016;15:1273-84.
5. Dandinoğlu T, Karadeniz M, Yılmaz V, Tekin L, Dinçer Ü. Evaluating neuropathic complaints by DN4 and LANSS scales after local corticosteroid therapy in carpal tunnel syndrome. J Back Musculoskelet Rehabil. 2016;29:575-80.

Letter to the Editor

DOI: 10.4274/haseki.galenos.2026.95866
Med Bull Haseki 2026;64(5):462-463

Beyond One-Hour Outcomes in Pediatric Gastroenteritis Rehydration

Emre Aygun

University of Health Sciences Türkiye, Sisli Hamidiye Etfal Training and Research Hospital, Clinic of Pediatrics, Istanbul, Türkiye

To the editor,

We read with interest the article by Yilmaz Donmez et al. (1) evaluating isotonic saline and Ringer's lactate for initial fluid resuscitation in children with acute gastroenteritis [Outcomes after initial rehydration with isotonic saline or Ringer's lactate in pediatric acute gastroenteritis (1-59 months): a single-center retrospective cross-sectional study]. The authors reported greater improvements in dehydration scores and in acid-base parameters with Ringer's lactate and suggested that it may be preferred as the initial replacement solution. We believe this study addresses an important clinical question; however, several aspects of the findings deserve further discussion.

The two groups differed in the severity of dehydration before treatment. Children in the Ringer's lactate group had higher dehydration scores at presentation. Because the Ringer's lactate group was more severely dehydrated at baseline, a greater absolute improvement after treatment may simply reflect the higher initial dehydration scores. Therefore, this baseline difference should be considered when interpreting the reduction in dehydration scores after treatment. In addition, because of the retrospective study design, physicians may have selected the intravenous fluid according to the patient's initial clinical severity, introducing potential treatment-selection bias. These factors make it difficult to determine how much of the observed improvement was attributable to the fluid itself. Recovery was assessed only after the first hour of treatment. Important clinical outcomes such as hospital stay, additional fluid requirements, treatment failure, PICU admission, and readmission were not evaluated. Previous studies have shown that early biochemical improvement does not always reflect overall clinical recovery (2).

Better laboratory values do not always mean better clinical outcomes. In a randomized trial, children treated with Ringer's lactate and isotonic saline had similar clinical recovery and hospital stays (2). A recent Cochrane review found better acid-base parameters with balanced solutions, but the clinical benefits were less clear (3). Although European Society of Paediatric and Neonatal Intensive Care guidelines support the use of balanced solutions, the supporting evidence is still limited (4).

The results of this study are clinically relevant. However, the apparent advantage of Ringer's lactate should be interpreted carefully. Although balanced solutions may improve acid-base parameters, current evidence has not consistently shown better clinical outcomes than isotonic saline. Prospective randomized studies powered for patient-centered clinical outcomes are needed before balanced crystalloids can be considered superior to isotonic saline.

Footnotes

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

References

1. Yilmaz Donmez C, Okay B, Hatipoglu HU, Sahin K. Outcomes after initial rehydration with isotonic saline or ringer's lactate in pediatric acute gastroenteritis (1-59 months): a single-center retrospective cross-sectional study. Med Bull Haseki. 2025;63:258-66.

Corresponding Author: Emre Aygun, MD, University of Health Sciences Türkiye, Sisli Hamidiye Etfal Training and Research Hospital, Clinic of Pediatrics, Istanbul, Türkiye
E-mail: emreaygun88@gmail.com **ORCID:** orcid.org/0000-0002-0546-327X
Received: 14.06.2026 **Accepted:** 17.07.2026 **Publication Date:** 17.11.2026

Cite this article as: Aygün E. Beyond one-hour outcomes in pediatric gastroenteritis rehydration. Med Bull Haseki. 2026;64(5):462-463

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

2. Kartha GB, Rameshkumar R, Mahadevan S. Randomized double-blind trial of ringer lactate versus normal saline in pediatric acute severe diarrheal dehydration. J Pediatr Gastroenterol Nutr. 2017;65:621-6.
3. Florez ID, Sierra J, Pérez-Gaxiola G. Balanced crystalloid solutions versus 0.9% saline for treating acute diarrhoea and severe dehydration in children. Cochrane Database Syst Rev. 2023;5:CD013640.
4. Brossier DW, Tume LN, Briant AR, et al. ESPNIC clinical practice guidelines: intravenous maintenance fluid therapy in acute and critically ill children- a systematic review and meta-analysis. Intensive Care Med. 2022;48:1691-708.

Corrigendum

DOI: 10.4274/haseki.galenos.2026.c001
Med Bull Haseki 2026;64(5):464

DOI: 10.4274/haseki.galenos.2026.63634
Med Bull Haseki 2026;64(4):244-253

DOI: 10.4274/haseki.galenos.2026.63634

In the article entitled "Diagnostic Accuracy of the Dietary Screening Tool for Detecting Malnutrition in Hospitalized Older Adults: Comparison with the Mini Nutritional Assessment-1999 and Nutritional Risk Screening-2002" (Med Bull Haseki. 2026;64(4):244-253; DOI: 10.4274/haseki.galenos.2026.63634), references 36 and 37 were incorrectly cited on page 250 due to an error made by the authors. At the authors' request, these references have been removed from the relevant sentence. The correction is shown in bold below.

The uncorrected version is as follows:

The relatively low sensitivity but high specificity of the DST obs (36,37).

The corrected version is as follows:

The relatively low sensitivity but high specificity of the DST obs.

©Copyright 2026 The Author(s). Published by Galenos Publishing House on behalf of Istanbul Haseki Training and Research Hospital.
This is an open access article under the Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 (CC BY-NC-ND) International License.

2026 Reviewer Index

Abdullah Agin
Abdullah Sakin
Abdulrahman Ozel
Ahmet Aydın
Alev Kural
Aliye Celikkol
Arda Isik
Atilla Bulur
Aysenur Erbin
Beliz Bilgili
Bilal Cetin
Egemen Cebeci
Emre Ozmen
Eray Kurnaz
Fatma Sarac
Gamze Citlak
Gamze Kucukosman
Genco Gencdal
Gulhan Tunca Sahin
Gunce Basarir
Hamide Piskinpasa
Hanife Serife Aktas
Hasim Olgun
Hayriye Esra Ataoglu
Ihsan Esen
Incinur Genisol
Isil Bakirci
Ismail Selvi
Kahraman Yakut
Levent Verim
Mahmut Civilibal
Medine Alpdemir
Mehmet Akif Kilic
Mehmet Senes
Menguhan Araz Altay
Mert Zure
Merve Damla Korkmaz
Muhammet Hamza Halil Toprak
Mustafa Zanyar Akkuzu
Nihan Hande Akcakaya
Nuri Alper Sahbaz
Omer Faruk Calim
Ozden Aksu
Ozgur Tanriverdi
Ramazan Yilmaz
Resit Yigit Yilancioglu
Gokcen Gozubatik Celik
Selami Aykut Temiz
Serhan Yilmaz
Sevgi Kulakli
Sezai Ozkan
Sibel Karsidag
Soner Sertan Kara
Taliha Oner
Taylan Onat
Tolgahan Atci
Tugba Sahbaz
Ugur Topal
Vedat Bakan
Veli Avci
Yasar Gokhan Gul
Zeynep Karaali
Zeynep Koc