E-ISSN: 2547-9431

BAGCILAR MEDICAL BULLETIN

Bağcılar Tıp Bülteni

Volume 11, Issue 2, June 2026

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

BAGCILAR MEDICAL BULLETIN
Bağcılar Tıp Bülteni

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

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

ADVISORY BOARD / DANIŞMA KURULU

ADVISORY BOARD / DANIŞMA KURULU

Ramazan Konkan
University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Clinic of Psychiatry, İstanbul, Turkey

Radiology

Mehmet Öncü
University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Clinic of Radiology, İstanbul, Turkey
Fethi Emre Ustabaşıoğlu
Trakya University Faculty of Medicine, Department of Radiology, Tekirdağ, Turkey
Serap Baş
Medipol Bahçelievler Hospital, Clinic of Radiology, İstanbul, Turkey

Radiation Oncology

Şule Karabulut Gül
University of Health Sciences Turkey, Kartal Dr. Lütfi Kırdar City Hospital, Clinic of Radiation Oncology, İstanbul, Turkey

Rheumatology

Selda Çelik
University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Clinic of Rheumatology, İstanbul, Turkey

Urology

Atilla Semerciöz
University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Clinic of Urology, İstanbul, Turkey
Emre Karabay
University of Health Sciences Turkey, Haydarpaşa Numune Training and Research Hospital, Clinic of Urology, İstanbul, Turkey

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

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

The journal is published online.

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

Responsible Manager: Assoc. Prof. Murat Altuntaş

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

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

Dergi, çevrimiçi olarak yayımlanmaktadır.

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

Sorumlu Yönetici: Doç. Dr. Murat Altuntaş

BAGCILAR MEDICAL BULLETIN
Bağcılar Tıp Bülteni

CONTENTS / İÇİNDEKİLER

REVIEWS / DERLEMELER

ORIGINAL RESEARCHES / ARAŞTIRMALAR

BAGCILAR MEDICAL BULLETIN
Bağcılar Tıp Bülteni

CONTENTS / İÇİNDEKİLER

REVIEW

Bagcilar Med Bull 2026;11(2):160-163
DOI: 10.4274/BMB.galenos.2025.29494

Probiotics for the Prevention of Sepsis in Preterm Infants: A Comprehensive Review of Recent Evidence

Preterm Bebeklerde Sepsisin Önlenmesinde Probiyotiklerin Rolü: Güncel Kanıtlara Dayalı Kapsamlı Bir Derleme

Emrah Can, Şahin Hamilçıkan, Yakup Karakurt, Cevat Yıldırım

University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Clinic of Neonatal Intensive Care, İstanbul, Turkey

Abstract

Preterm infants are highly susceptible to late-onset sepsis due to immaturity of their immune system and imbalances in gut microbiota. In recent years, probiotics have emerged as a promising strategy to reduce the incidence of necrotizing enterocolitis (NEC) and sepsis in this vulnerable group. This review synthesizes findings from randomized controlled trials and meta-analyses published between 2020 and 2025, focusing on the effects of probiotic supplementation on NEC, sepsis, and overall mortality. Studies were identified through systematic searches of PubMed, Cochrane, and Scopus databases. The analysis also considers the most commonly studied strains, dosing regimens, timing of administration, safety data, and current clinical recommendations. Across the literature, probiotic use is consistently associated with a marked reduction in NEC, and some evidence indicates modest reductions in mortality. However, the impact on sepsis prevention varies considerably depending on strain combinations, trial design, and clinical context. While cases of probiotic-related sepsis are rare, ensuring product safety and quality control remains essential. When appropriate strains are used under careful clinical oversight, probiotics represent a valuable adjunctive strategy in the care of preterm infants, particularly for NEC prevention.

Keywords: Gut microbiota, necrotizing enterocolitis, neonate, preterm infants, probiotics, sepsis

Öz

Prematüre bebekler, immün sistemlerinin henüz olgunlaşmamış olması ve bağırsak mikrobiyotasındaki dengesizlik nedeniyle geç başlangıçlı sepsis ve nekrotizan enterokolit (NEC) açısından yüksek risk altındadır. Son yıllarda probiyotikler, bu komplikasyonların önlenmesinde destekleyici bir yaklaşım olarak dikkat çekmektedir. Bu derlemede, 2020-2025 yılları arasında yayımlanan randomize kontrollü çalışmalar ve meta-analizler incelenerek, probiyotik desteğinin NEC, sepsis ve mortalite üzerindeki etkileri değerlendirilmiştir. Literatür taraması PubMed, Cochrane ve Scopus veri tabanları kullanılarak sistematik olarak yapılmıştır. Derlemede ayrıca en sık kullanılan suşlar, doz rejimleri, uygulama zamanı, güvenlik verileri ve uluslararası kılavuz önerileri ele alınmıştır. Bulgular, özellikle çoklu suş içeren preparatların NEC insidansını anlamlı düzeyde azalttığını ve mortalite oranlarında da sınırlı da olsa olumlu etkiler sağlayabildiğini göstermektedir. Sepsisi önleyici etkiler ise çalışmalara göre değişkenlik göstermekte; kullanılan suşlara, çalışmanın kalitesine ve klinik koşullara bağlı olarak farklılık arz etmektedir. Probiyotik kaynaklı sepsis olguları oldukça nadir bildirilmiş olsa da, kullanılan ürünlerin kalite kontrolü ve güvenilirliği kritik önemdedir. Uygun suşlar dikkatli bir klinik gözetim altında kullanıldığında, probiyotikler özellikle NEC'nin önlenmesinde değerli bir destekleyici yaklaşım olarak öne çıkmaktadır.

Anahtar kelimeler: Mikrobiyota, nekrotizan enterokolit, preterm bebek, probiyotik, sepsis, yenidoğan

Address for Correspondence: Emrah Can, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Clinic of Neonatal Intensive Care, İstanbul, Turkey
E-mail: canemrahcan@yahoo.com **ORCID:** orcid.org/0000-0002-2163-6564
Received: 15.04.2025 **Accepted:** 18.11.2025 **Epub:** 20.11.2025 **Publication Date:** 24.06.2026

Cite this article as: Can E, Hamilçıkan Ş, Karakurt Y, Yıldırım C. Probiotics for the prevention of sepsis in preterm infants: a comprehensive review of recent evidence. Bagcilar Med Bull. 2026;11(2):160-163

Introduction

Preterm infants, especially those born before 32 weeks of gestation, are highly vulnerable to severe infections due to their immature immune systems and the frequent need for invasive medical interventions (1). Late-onset sepsis (LOS), typically diagnosed after the third day of life, remains a major threat to both survival and clinical stability in this population (2). Necrotizing enterocolitis (NEC), another life-threatening complication, often precedes or coexists with LOS and is characterized by systemic inflammation and intestinal injury (3). Disruptions in the gut microbiota of preterm neonates, characterized by insufficient colonization by beneficial microbes and overgrowth of pathogens, play a central role in the development of both NEC and LOS (4).

Probiotics, defined as live microorganisms that confer health benefits when administered in adequate amounts, have been proposed as a preventive strategy in this context. They may help restore microbial balance, strengthen the intestinal barrier, and modulate immune responses. Numerous randomized controlled trials and meta-analyses have investigated the role of probiotics in reducing the incidence of NEC, sepsis, and mortality in preterm neonates.

This review brings together current evidence on the use of probiotics in preterm infants, with a focus on clinical outcomes, safety data, commonly studied strains, dosing strategies, and recommendations from key health authorities.

Methods

We conducted a structured literature search using three major biomedical databases: PubMed, Scopus, and the Cochrane Library. The search covered publications from January 2020 to March 2025 and focused on the following search terms: Probiotics, preterm, sepsis, NEC, and neonates. The selection process prioritized randomized controlled trials, systematic reviews, and meta-analyses that assessed the effects of probiotics on clinical outcomes in preterm infants. Studies were included based on their relevance to the topic, methodological quality, and reporting of key outcomes such as NEC, LOS, mortality, and adverse events. The findings were synthesized narratively to provide a comprehensive overview of current evidence.

Efficacy of Probiotics on Sepsis Prevention

Evidence regarding the efficacy of probiotics in reducing LOS in preterm infants remains inconclusive. A 2023 Cochrane meta-analysis update encompassing more than 11,000 infants confirmed that probiotics significantly reduced the incidence of NEC and may reduce mortality, whereas their effect on LOS was minimal or uncertain (5).

Conversely, a Bayesian network meta-analysis by Thomas et al. (6), analyzing 29 trials, reported that a specific combination of four probiotic strains (*B. longum, B. bifidum, B. infantis*, and *L. acidophilus*) significantly lowered sepsis incidence. Nevertheless, the overall certainty of the evidence was limited because of heterogeneity in trial design and regional variation (6). Another comprehensive network meta-analysis published in JAMA Pediatrics in 2023, which included more than 25,000 preterm infants from 106 RCTs, concluded that combinations of probiotics and lactoferrin were more effective in reducing sepsis rates than probiotics alone (7).

Recent large-scale reviews underscore that while the benefit of NEC reduction is well established, sepsis outcomes are highly variable and appear to depend on the specific strain used, dosage, and patient population. For instance, Deshpande et al. (8), in a review of over 70,000 infants, observed significant reductions in NEC and mortality but reported inconsistent findings regarding LOS.

Probiotic Effects on NEC and Mortality

The most consistent finding in the literature is a significant reduction in NEC incidence following probiotic administration. Meta-analyses have shown that probiotic use can reduce NEC (Bell stage ≥II) by approximately 40 to 60 percent, especially when multi-strain formulations are employed (5,7). This protective effect is largely attributed to improved intestinal barrier function and suppression of pathogenic bacteria.

Since NEC is a known precursor to secondary sepsis and mortality, preventing NEC may indirectly reduce sepsis-related deaths. Some studies have reported reductions of 25-30 percent in all-cause neonatal mortality with probiotic use (7,8). However, the extent to which these mortality benefits are directly linked to sepsis prevention remains unclear.

Probiotic Strains, Doses and Administration

The most widely studied probiotic strains in preterm infants include *Lactobacillus rhamnosus* GG, *L. reuteri*, *Bifidobacterium breve, B. longum*, and *B. lactis*. Multi-strain preparations tend to offer broader benefits than single-strain formulations. Effective regimens generally contain between 10^8 and 10^9 colony-forming units per strain per

day. Administration typically begins within the first few days of life and continues until 34-36 weeks' postmenstrual age (6,8,9).

The 2020 ESPGHAN position paper identified two probiotic approaches with the strongest supporting evidence. These include *L. rhamnosus* GG alone and a combination of *B. infantis* Bb-02, *B. lactis* Bb-12, and *Streptococcus thermophilus* TH-4 (10). Updated guidance emphasizes the importance of preparation quality, dose standardization, and rigorous monitoring to ensure safety and efficacy.

Safety and Probiotic-associated Sepsis

Overall, probiotics appear to be well tolerated in preterm infants. Randomized trials and observational studies report similar rates of adverse events between probiotic and control groups. A 2022 comprehensive review documented only 32 cases of probiotic-associated sepsis, mostly involving species of Bifidobacterium or *Lactobacillus*; the majority of affected infants recovered fully after antibiotic therapy (11).

Despite the low incidence, safety concerns persist. The United States Food and Drug Administration has issued advisories cautioning against the routine use of probiotics in preterm neonates due to risks associated with contamination and variable product quality (12). A fatal case of fungal sepsis linked to a contaminated product has heightened awareness of the need for pharmaceutical-grade formulations and stringent manufacturing standards.

Clinical Guidelines and Global Practices

Clinical recommendations regarding probiotic use in preterm infants vary across regions. The American Academy of Pediatrics does not support routine probiotic use due to unresolved safety concerns, particularly in extremely low birth weight infants (13). In contrast, ESPGHAN and the World Health Organization conditionally recommend selected probiotic strains for preterm infants fed with human milk, provided that high-quality products are used (10,14).

In some countries, including the United Kingdom and Canada, national neonatal networks have integrated routine probiotic supplementation into NICU protocols based on favorable risk-benefit profiles.

As concerns persist about live microbial supplementation, attention is shifting toward non-viable alternatives such as paraprobiotics. These inactivated microbial preparations maintain immunomodulatory properties while minimizing the risk of systemic infection. Emerging evidence from preclinical studies suggests that paraprobiotics may strengthen gut barrier integrity and reduce inflammation in models of NEC and colitis. For example, Batista et al. (15) demonstrated that paraprobiotics derived from *Lactobacillus delbrueckii* CIDCA 133 alleviated 5-FU-induced intestinal inflammation in neonatal mice. Sundram et al. (16) and Li et al. (17) provided further support for the protective potential of heat-inactivated *Lactiplantibacillus plantarum* strains against microbial dysbiosis and gut injury. However, clinical trials in preterm infants are currently limited, and further investigation is warranted.

A recent large-scale cohort study conducted in Canadian NICUs reported that probiotic use in preterm infants was associated with a significant reduction in mortality but not with a significant decrease in the incidence of NEC or LOS (18).

According to the 2025 guidelines of the Turkish National Neonatology Association, the routine use of probiotics is not recommended in enteral feeding protocols for preterm infants. However, if the unit chooses to administer probiotics and has the necessary conditions in place, specific strains with proven safety and efficacy may be considered for infants born before 32 weeks' gestation and weighing between 1.000 and 1.500 grams. In these cases, clinicians are advised to inform families about potential benefits and risks and to obtain informed consent. Routine use in infants with a birth weight below 1.000 grams is not currently recommended (19).

Conclusion

Probiotic supplementation in preterm infants consistently reduces the incidence of NEC and may contribute to a modest decrease in overall mortality. However, its direct impact on sepsis prevention remains unclear. While some combinations, especially those that include lactoferrin, show promise, inconsistencies in study findings and heterogeneity in probiotic formulations limit the ability to draw broad conclusions. Although cases of probiotic-associated sepsis are rare, careful monitoring and strict quality control are essential. Future research should focus on defining optimal strains, dosages, and duration of therapy and ensuring pharmaceutical-grade production. As evidence continues to accumulate, probiotics hold promise as a supportive intervention in neonatal care, provided their use is guided by high-quality data and individualized clinical judgment.

Footnotes

Authorship Contributions

Concept: E.C., Ş.H., Y.K., C.Y., Design: E.C., Ş.H., Y.K., C.Y., Data Collection or Processing: Y.K., C.Y., Analysis or Interpretation: E.C., Ş.H., Literature Search: Y.K., C.Y., Writing: E.C., Ş.H., Y.K., C.Y.

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

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

References

1. Ma Y, Peng X, Zhang J, Zhu Y, Huang R, Li G, et al. Gut microbiota in preterm infants with late-onset sepsis and pneumonia: a pilot case-control study. BMC Microbiol. 2024;24(1):272.
2. Shane AL, Sánchez PJ, Stoll BJ. Neonatal sepsis. Lancet. 2017;390(10104):1770-1780.
3. Niño DF, Sodhi CP, Hackam DJ. Necrotizing enterocolitis: new insights into pathogenesis and mechanisms. Nat Rev Gastroenterol Hepatol. 2016;13(10):590-600.
4. Kim CS, Claud EC. Necrotizing enterocolitis pathophysiology: how microbiome data alter our understanding. Clin Perinatol. 2019;46(1):29-38.
5. Sharif S, Meader N, Oddie SJ, Rojas-Reyes MX, McGuire W. Probiotics to prevent necrotising enterocolitis in very preterm or very low birth weight infants. Cochrane Database Syst Rev. 2023;7(7):CD005496.
6. Thomas JP, Raine T, Reddy S, Belteki G. Probiotics for the prevention of necrotising enterocolitis in very low-birth-weight infants: a meta-analysis and systematic review. Acta Paediatr. 2017;106(11):1729-1741.
7. Wang Y, Florez ID, Morgan RL, Foroutan F, Chang Y, Crandon HN, et al. Probiotics, prebiotics, lactoferrin, and combination products for prevention of mortality and morbidity in preterm infants: a systematic review and network meta-analysis. JAMA Pediatr. 2023;177(11):1158-1167. Erratum in: JAMA Pediatr. 2024;178(1):99.
8. Deshpande G, Jape G, Rao S, Patole S. Benefits of probiotics in preterm neonates in low-income and medium-income countries: a systematic review of randomised controlled trials. BMJ Open. 2017;7(12):e017638.
9. Athalye-Jape G, Rao S, Patole S. Lactobacillus reuteri DSM 17938 as a probiotic for preterm neonates: a strain-specific systematic review. JPEN J Parenter Enteral Nutr. 2016;40(6):783-794.
10. van den Akker CHP, van Goudoever JB, Shamir R, Domellöf M, Embleton ND, Hojsak I, et al. Probiotics and preterm infants: a position paper by the European Society for Paediatric Gastroenterology Hepatology and Nutrition Committee on Nutrition and the European Society for Paediatric Gastroenterology Hepatology and Nutrition Working Group for probiotics and prebiotics. J Pediatr Gastroenterol Nutr. 2020;70(5):664-680.
11. Liu H, Wang B, Lu T, Pei Y. Safety and efficacy of probiotics in the prevention of necrotizing enterocolitis in premature and/or low-birthweight infants: a systematic review and meta-analysis. Transl Pediatr. 2022;11(2):249-259.
12. US Food and Drug Administration (FDA). Important safety information: FDA advises hospitals to avoid using probiotic products in preterm infants. [Internet]. 2023. Available from:https://www.fda.gov/food/alerts-advisories-safety-information
13. Poindexter B; Committee on Fetus and Newborn. Use of probiotics in preterm infants. Pediatrics. 2021;147(6):e2021051485. Erratum in: Pediatrics. 2021;148(6):e2021054370.
14. Care of Preterm or Low Birthweight Infants Group. New World Health Organization recommendations for care of preterm or low birth weight infants: health policy. EClinicalMedicine. 2023;63:102155.
15. Batista VL, De Jesus LCL, Tavares LM, Barroso FLA, Fernandes LJDS, Freitas ADS, et al. Paraprobiotics and postbiotics of *Lactobacillus delbrueckii* CIDCA 133 mitigate 5-FU-induced intestinal inflammation. Microorganisms. 2022;10(7):1418.
16. Sundram TKM, Tan ESS, Lim HS, Amini F, Bustami NA, Tan PY, et al. Effects of ambient particulate matter ($PM_{2.5}$) exposure on calorie intake and appetite of outdoor workers. Nutrients. 2022;14(22):4858.
17. Li S, Tantibhadrasapa A, Buddhasiri S, Boonpan P, Sukjoi C, Mongkolkarvin P, et al. Probiotic, paraprobiotic, and postbiotic activities of *Lactiplantibacillus plantarum* KUNN19-2 against non-typhoidal *Salmonella* serovars. Int J Mol Sci. 2025;26(5):1821.
18. Alshaikh BN, Ting J, Lee S, Lemyre B, Wong J, Afifi J, et al. Effectiveness and risks of probiotics in preterm infants. Pediatrics. 2025;155(3):e2024069102.
19. The guideline of the enteral feeding of preterm neonates. [Internet]. Turkish Neonatal Society Guidelines 2025. [cites 2025 Apr 13];46. Available from:https://neonatology.org.tr/tani-tedavi-protokolleri

REVIEW

Bagcilar Med Bull 2026;11(2):164-176
DOI: 10.4274/BMB.galenos.2026.15807

Artificial Intelligence, Representation, Language Models, Cognitive Science, and Applications in Psychiatry

Yapay Zeka, Temsiliyet, Dil Modelleri, Bilişsel Bilim ve Psikiyatride Uygulamaları

Fırat Belli[1], Ali Rıza Türkmen[2]

[1]İstanbul University, Department of Linguistics, İstanbul, Turkey
[2]University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Psychiatry, İstanbul, Turkey

Abstract

The paper explores the issue of representation in artificial intelligence (AI) by linking technical models (symbolic, connectionist/hybrid) with Wittgenstein's philosophy of language and delineates implications for cognitive science and psychiatric practice. The issue of representation remains at the core of developing trustworthy, interpretable systems and is closely tied with what is defined by AI—systems rationally behaving, human-like reasoners, or human-like actors. Representation formalisms hold, manipulate, and interpret knowledge in specific modes; shifting towards deeper learning and transformer constructions increased large language model's ability in processing natural language. Wittgenstein's picture theory and his language games open a window into how meaning is constructed in such models, making explainable AI imperative for transparency as much as for gaining trust. From a cognitive-scientific perspective, the key issue is how accurately AI can approximate human cognition. Skills such as multitask or few-shot learning or contextual inference are similarities between human learning abilities, but no current systems are in a position where they can interpret cultural context, "forms of life," or multimodal meaning. Future quantum computing, integration of neuroscientific signals (e.g., electroencephalogram), and multimodal pipelines would possibly redesign representational frameworks and generate contextually more sensitive systems. For psychiatry, symbolic, connectionist, and hybrid methods already include applications such as depression detection, risk prediction for schizophrenia, and enhanced social interaction for individuals with

Öz

Bu makale, teknik modelleri (sembolik, bağlantısal, hibrit) Wittgenstein'ın dil felsefesi ile ilişkilendirerek yapay zekadaki (YZ) temsil sorununu incelemeyi ve bunların bilişsel bilim ve psikiyatri pratiği için çıkarımlarını vurgulamayı amaçlamaktadır. YZ'de temsil sorunu, güvenilir ve yorumlanabilir sistemler geliştirmede karşılaşılan temel zorluklardan biridir. Bu sorun, YZ'nin tanımıyla yakından ilişkilidir ve rasyonel davranış, insan benzeri düşünme ve insan benzeri davranış yaklaşımları üzerinden ele alınabilir. Sembolik, bağlantısal ve hibrit modeller gibi temsil yöntemleri, bilginin kodlanması, işlenmesi ve yorumlanması için farklı mekanizmalar sunar. İstatistiksel yaklaşımlardan derin öğrenme ve transformer tabanlı mimarilere geçiş, büyük dil modellerinin doğal dili işleme yeteneklerinde önemli ilerlemeler sağlamıştır. Wittgenstein'ın "resim kuramı" ve "dil oyunları" yaklaşımları, bu modellerde anlam inşasını anlamak için kavramsal bir çerçeve sunar. Açıklanabilir YZ, şeffaflık ve güvenin artırılması açısından kritik bir yaklaşım olarak öne çıkmaktadır. Bilişsel bilim perspektifinden, YZ'nin insan bilişsel süreçlerini ne ölçüde taklit edebileceği konusu önemli tartışmalardan biridir. Çoklu görev öğrenme, az örnekle öğrenme ve bağlamsal anlam çıkarma gibi yetenekler, insan zihninin öğrenme biçimleriyle benzerlik gösterse de, YZ hala kültürel bağlamı, yaşam formunu ve çoklu modalite üzerinden anlamı tam olarak yakalamakta sınırlıdır. Gelecekte kuantum bilişim, sinirbilim verilerinin (örn, elektroensefalogram) entegrasyonu ve çoklu modalite sistemleri, temsil stratejilerini yeniden şekillendirme ve

Address for Correspondence: Ali Rıza Türkmen, MD, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Psychiatry, İstanbul, Turkey
E-mail: drturkmen20@gmail.com **ORCID:** orcid.org/0009-0007-3756-698X
Received: 28.08.2025 **Accepted:** 15.04.2026 **Epub:** 29.04.2026 **Publication Date:** 24.06.2026

Cite this article as: Belli F, Türkmen AR. Artificial intelligence, representation, language models, cognitive science, and applications in psychiatry. Bagcilar Med Bull. 2026;11(2):164-176

Abstract

dementia. Advancements for psychotherapy-directed descriptions of language can further raise utility. Integrating philosophical, technical, and clinical perspectives on representation can guide design for more interpretable, safe AI with clinical utility in psychiatry.

Keywords: Artificial intelligence, cognitive science, large language models, psychiatry, representation, Wittgenstein

Öz

YZ'yi bilişsel bilim ile bütünleştirerek daha insan benzeri, bağlamı yakalayabilen yetenekler kazandırma potansiyeline sahiptir. Bununla birlikte, psikiyatri uygulamalarında kullanılan sembolik, bağlantısal ve hibrit yaklaşımlar oldukça önemlidir. YZ'nin depresyon tespiti, şizofreni riskinin öngörülmesi, demans bakımında sosyal etkileşimin artırılması gibi çok çeşitli klinik uygulamaları mevcuttur. Psikoterapi ve dil temsili alanındaki gelişmelerin gelecekte YZ'nin psikiyatrideki işlevselliğini artıracağı öngörülmektedir. Genel olarak, temsil sorununa ilişkin felsefi, teknik ve klinik bakış açılarının bütünleştirilmesi, psikiyatride daha yorumlanabilir, daha güvenli ve klinik açıdan daha yararlı YZ sistemlerine giden bir yol haritası sunar.

Anahtar kelimeler: Bilişsel bilim, büyük dil modelleri, psikiyatri, temsil, Wittgenstein, yapay zeka

Introduction

A review of the literature reveals that artificial intelligence (AI) has persisted as a field that has intensively utilized engineering, linguistics, cognitive science, philosophy, and neuroscience since the mid-twentieth century. Alan Turing's question in 1950, which surprised most people, "Can machines think?", has led to interesting and surprising results in people's minds. Turing's question marked one of the earliest conceptual turning points in AI research sparking lasting debates on machine intelligence and cognition (1). The rise of computer science and technology has been highly significant. As algorithms began mimicking how our minds and language work, AI evolved from just a technical field to one that looks at modeling key aspects of human thought like language, reasoning, memory, and decision-making. Over time, various ideas about AI emerged, shaped by studies in language and philosophy, alongside rapid technological advancements. In this context, the concept of representation how AI systems gather, organize, and apply knowledge about the world has been explored in depth.

Language ability is one of the most important capacities that distinguish humans from other primates. Language also constitutes one of the most fundamental cognitive abilities of the human mind and plays a central role in thinking, communication, information sharing, and social interaction. Such an important human ability cannot be addressed in isolation from discussions about its representation in AI. Analyzing language solely as a means of expression would be insufficient. Language can also be described as an element that embodies the semantic and representational dimensions of objects and events in the world. For AI systems to resemble human cognitive abilities, or at least mimic them, they must incorporate language-based representation mechanisms. Accordingly, language processing in AI research is not only a technical challenge but also an area where philosophical and linguistic questions related to meaning, context, and cultural intertwining arise. AI architectures predominantly use language models. These language models are designed to process, interpret, and generate human language. Recently developed large language models (LLMs) have significantly advanced natural language processing (NLP) by demonstrating high performance on language generation and comprehension tasks (2). LLMs are now important tools that most people are aware of. They have permeated almost every aspect of our lives. These discussions and developments necessarily entail philosophical considerations. Philosophy, and especially philosophy of language, is inevitably part of the process. From the perspective of the philosophy of language, these developments must focus on theories of language, particularly the relationships among language use, meaning, and context. In this context, it is inevitable to revisit Ludwig Wittgenstein's approaches.

Recently, the latest deep learning approaches built on NLP have become quite popular. These approaches are highly functional and efficient. They are logic-based and rule-oriented models, and their importance in this field is growing due to their transparency and interpretability (3). Language ability is seriously related to logical representation. Of course, emphasizing this relationship alone is not sufficient. But representing language with formal logic is very important. This representation contributes significantly to revealing the abilities of explicit semantic inference and structured reasoning, which are still challenging for data-driven systems. Historically, AI research has encompassed several fundamental approaches to language processing.

Generally, these consist of N-gram models, recurrent neural networks (RNN), and transformer-based architectures (4). All of these approaches are used in the creation of today's AI models. In particular, the models include data-driven methods based on statistical learning from large data sets. They also include rule-based mechanisms that process symbolic and logical representations of meaning, which is important for our topic. Each approach offers different advantages and limitations, particularly in terms of semantic precision and contextual understanding. NLP, initially developed as one of AI's most dynamic subfields, was built using linguistic theory and cognitive models aiming to simulate human language comprehension. Over time, especially with the increasing effectiveness of deep learning methods, this field has shifted towards a paradigm focused on technological developments aiming for high- performance expectations. LLMs trained on large datasets have achieved remarkable success in tasks such as text generation, translation, summarization, dialogue, and question answering. However, whether such systems truly capture semantic understanding or primarily reflect statistical regularities inherent in language use remains a subject of debate (3,5). They can't capture human cognition and language capacity at this level. Perhaps we will reach a more perfect point in the future. But for now, it does not seem possible to express this. In fact, this article focuses on how this goal can be partially achieved in some areas. In addition to these developments, AI is increasingly able to mimic human cognitive abilities in areas such as memory, attention, problem-solving, and decision-making. Given these capabilities, it has the potential to deepen and enhance existing human capacity (6). In light of these developments, particular emphasis should be placed on the mutual interaction between AI and cognitive science. Advances in the study of human cognition provide information for the design of AI models, while AI systems also broaden research opportunities by bringing new perspectives to cognitive processes. This represents a two-way process of development and mutual influence. Such interaction is particularly important in areas where language-based reasoning and interpretation are central. Neuroscience research can also be conducted more easily using AI tools. All of these capabilities can help put significant opportunities at the service of science.

AI applications are widely used across the social, basic, and medical sciences. Among these fields, medical applications and, more specifically, areas such as psychiatry form part of our discussion. In these fields, language, cognition, and meaning play a decisive role. In light of this, careful attention should be paid to the advancement of AI applications in psychiatry. Psychiatric practice and theoretical debates, patient assessment and follow-up, and therapeutic assistance have all benefited from the deployment of AI applications in recent years (7). These developments have been amplified by the coronavirus disease-2019 (COVID-19) pandemic. The demand for mental health treatments was greatly raised by this epidemic, but clinical capacity was similarly constrained. Thus, the relationship between AI and psychiatry underwent a sea change as a result of the epidemic. Access to this service has become even more crucial due to declining interpersonal interactions, rising suicide rates from mental illnesses, social isolation, and increased substance use (8). From a broader perspective, the COVID-19 crisis has increased global interest in AI technologies and accelerated progress across many areas.

The topics briefly summarized above are not independent of each other. On the contrary, they are deeply interconnected. The primary aim of this article is to subject this interconnection to a critical examination as comprehensively as possible. In this context, the article aims to discuss how AI, particularly language-based models, is conceptualized and applied in terms of representation, cognition, and language. Assessing the implications of these developments for medical and psychiatric education is also crucial. First, this analysis will address debates concerning logical and statistical representations in NLP. It will then examine the relationships between AI, linguistics, representation, philosophy, and cognitive models, while also discussing the extent to which AI systems can support or enhance human cognitive functions. Ultimately, framing these discussions within the context of medical and psychiatric practice will enrich the analysis and clarify its clinical relevance. This article will argue that representation, meaning, and context are extremely important in coding human language and other cognitive abilities in AI applications. These applications may be important in producing more efficient and human- like outputs in theoretical and practical fields.

Definitions of AI

First, it is necessary to establish a conceptual framework regarding the definition of AI. Although various definitions of AI have been proposed since the 1950s, it is important to clarify this issue to some extent. AI can be defined in various ways depending on the theoretical and methodological perspective adopted. The literature reflects a wide variety of definitions, consistent with the interdisciplinary nature of this field. Some authors divide these definitions into four

main categories: Systems that think like humans, systems that behave like humans, systems that think rationally, and systems that behave rationally (9). We must note from the outset that even these categorizations are insufficient. Since conceptual ambiguity complicates both theoretical discussion and practical evaluation, establishing a clear definition framework should be a primary goal. In this section, three effective approaches to rational action, human- like thinking, and human-like behavior are examined, as they are particularly relevant to discussions concerning cognition, language, and medical practice. These definitions will provide a suitable foundation for subsequent explanations and analyses.

Focused on Rational Action (Logical Approach)

The rational action approach defines AI as systems capable of selecting actions that maximize the probability of achieving predefined goals. Within this framework, intelligence is associated with the ability to make decisions up to a certain level under specific constraints. Of course, this decision-making also progresses through algorithms resembling language and cognitive abilities. A classic example would be a chess program that can evaluate possible moves in advance and select the one with the highest probability of achieving its goals. This approach relies heavily on formal knowledge representation, logical inference, and structured reasoning mechanisms. Here, representation and human-like abilities come into play. Historically, symbolic AI emerged from this tradition and has featured explicit rules and structured representations of knowledge (10). The strengths of this approach include transparency and logical consistency. However, its limitations become apparent in contexts characterized by uncertainty, ambiguity, or incomplete information. Therefore, it is difficult to define tools designed to represent the human mind solely in terms of these formal and rule-based characteristics. It is not even easy to mimic the human mind in depth.

Definition Focused on Thinking Like Humans (Cognitive Approach)

The cognitive approach conceptualizes AI as an attempt to model and replicate certain capacities of the human mind, including learning, reasoning, and problem-solving. This perspective represents one of the most significant intersections between cognitive science and AI research. Rather than focusing solely on observable behavior, it seeks to understand how humans process information and to translate these mechanisms into computational models. NLP systems exemplify this approach by attempting to simulate how humans comprehend, interpret, and generate linguistic responses (11). Although this perspective offers valuable insights into human cognition, accurately modeling complex mental processes remains a substantial scientific challenge. Nevertheless, it has become a highly influential and operationally effective framework within contemporary AI systems.

Definition Focused on Acting Like Humans (Pragmatic Approach)

The pragmatic approach evaluates AI primarily based on observable behavior rather than internal cognitive mechanisms. It should also be noted that this approach is based on linguistic production. According to this view, if a system's responses during interaction are indistinguishable from those of a human, it can be considered intelligent. Alan Turing's Turing Test is the most effective example of this approach and proposes behavioral indistinguishability as a criterion for intelligence (12). Of course, this behavioral similarity is not a motor output- like situation. It focuses on the nature of the written and spoken outputs produced. It can partially answer Turing's famous question. Today's LLMs are increasingly producing more natural and contextually appropriate interactions and therefore fall within the scope of this definition. However, some authors argue that human-like behavior does not necessarily imply true understanding or cognition. In this context, LLMs can produce outputs that simulate human cognitive capacities by utilizing large amounts of data. While producing such outputs, these systems may exhibit forms of expression that closely resemble human actions, but this similarity does not definitively prove the existence of genuine understanding. The development of LLMs is a crucial phase. The process is highly dynamic and fluid. It is not easy to predict where LLMs might evolve.

Representation in AI: Conceptual and Technical Dimensions

If knowledge, along with entities and even emotions belonging to the world, is not transformed into representational objects, it becomes exceedingly difficult to speak of any form of intelligence. Language itself, after all, consists fundamentally of a system of representations. We may now introduce a different perspective by turning to the issue of representation. Addressing representation will help clarify several key domains that will become central in the subsequent analyses.

In fact, the human mind contains a representation of the world. Language itself produces output through symbols and representations. In the context of AI, representation

refers to the meaningful and computable encoding of entities, events, or concepts from the real world within an artificial system (13). Similar to how a map selectively represents certain aspects of a city, AI-based representation involves translating the complexity of the world into structured formats that machines can process. This process is not merely technical; it also raises philosophical questions about how meaning is created and interpreted. According to Wittgenstein's later philosophy, meaning is not inherent in symbols but emerges from their use in specific contexts. In his view, "The limits of language are the limits of the world" (14). Consequently, the choice of representational framework directly affects what an AI system can represent, infer, and "understand". Historically, broadly speaking, AI research has yielded three main perspectives on representation.

Symbolic Representation (Symbolic AI)

We have discussed the capacity of language for symbolization and representation. To elaborate further, symbolic representation relies on the encoding of information through clearly defined symbols and formal logical rules. Its philosophical foundations are based on the views of many thinkers and philosophers throughout history. For example, it extends from Leibniz's view of a universal logic language to Boole's logic algebra and Frege's formalization of meaning. This approach has transformed into a different reality with Turing's theory of computation and McCarthy's (15) conceptual explanation of the term AI at the 1956 Dartmouth Conference. Technically, symbolic systems use structures such as production rules, semantic networks, various frame-based representations, and formal logical expressions. They are widely used in expert systems, logical programming, and certain forms of automated theorem proving. Their fundamental strengths are explainability and logical rigor. However, they face significant challenges in dealing with uncertainty, learning from data, and representing nuanced or context- dependent information (16). When context is lacking, an important aspect of reality is missing, and the human dimension awaits completion.

Connectionist Representation (Connectionist AI)

Connective representation positions information not as separate symbolic structures, but as distributed patterns across artificial neural networks. LLMs are largely an advanced combination of these approaches. Many authors have historically contributed to the connectionist representation approach. Among these fundamental contributions are Hebb's (17) relational learning approaches, Rosenblatt's (18) model, and the parallel distributed processing framework developed by Rumelhart and McClelland (19). In connectionist systems, information is encoded in weighted connections between units. Learning in this system occurs through the iterative adjustment of these weights during training. Words that describe objects are converted into semantic concepts and can be represented as high-dimensional vectors associated with multiple features. This paradigm underpins many contemporary applications, including image recognition, speech processing, and NLP systems such as bidirectional encoder representations from transformers (BERT) and generative pre- trained transformers (GPT). Connective models enable generalization and adaptive learning from examples, but their internal representations are often controversial and difficult to interpret (20). These models have quite broad areas of influence in today's world.

Hybrid Representation

Hybrid representation aims to combine the strengths of symbolic and connective approaches while reducing their limitations. Among the first examples of hybrid cognitive architectures are state, operator, and result (21) and adaptive control of thought-rational (22), which combine rule-based reasoning with learning mechanisms. In recent years, this perspective has been revisited in research on neuro-symbolic AI. In such systems, neural components perform perception and pattern recognition tasks, while symbolic components perform explicit reasoning and inference. This effectively combines the other two models. The transformer architecture, one of today's important tools, can also be considered a form of hybrid representation due to its capacity to integrate contextual learning with structured attention mechanisms (23). Hybrid systems offer a promising framework for achieving both high performance and interpretability, particularly in areas requiring reliable reasoning and explainable results. All of these models are used in many AI architectures. These explanations are written to reiterate how important concept representation is in AI models.

AI and Language Models: Technical Evolution and Wittgenstein's Perspective Evolution and Working Mechanisms of Language Models

After briefly summarizing the representation problem, we can now discuss language in general and, more specifically, language models in AI. In doing so, we will try to draw on philosophy a little. We have already made philosophical references here and there since the beginning of the article. However, this article is not based on a discussion of the philosophy of linguistics. But since we are trying to conduct

an interdisciplinary discussion here, we will not pass over it without comment.

Until the early 2010s, language modeling in AI tools was primarily based on statistical approaches. N-gram models, which could be considered outdated structures, predicted the probability of a word based on a fixed number of preceding words and could establish language connections in short texts, achieving reasonable success (24). However, these models were insufficient at maintaining semantic consistency in longer texts. In today's evolving world, there is a necessity to generate longer texts and perform complex tasks.

The development of deep learning techniques has been groundbreaking. These developments have greatly simplified tasks in language modeling. RNN and long short-term memory architectures enabled the system to capture sequential dependencies by better propagating information between text steps. Despite these developments, problems remained. Information decay in long sequences remained a problem. This decay kept effectiveness at a certain level in complex linguistic tasks (25). Overcoming this situation seemed inevitable. Human intellect and the technology it produced sought to overcome this as well.

Transformer architectures were subsequently developed. Finally, in 2017, a significant milestone was reached with the introduction of the Transformer architecture. Unlike sequential models, transformers use attention mechanisms that enable the simultaneous evaluation of relationships between all tokens in a sequence. This architectural innovation enabled more comprehensive and contextual modeling, significantly increasing capacity across a wide range of NLP tasks. Models built on this framework, such as OpenAI's GPT, BERT, LaMDA, and PaLM, have been trained on billions of parameters. Thus, LLMs have achieved widespread adoption across a broad range of domains. They perform at a high level not only in word prediction but also in areas such as translation, summarization, question answering, and code generation (4,26). These models have offered significant opportunities in fields such as medicine and psychiatry. AI models have reached this stage after accomplishing challenging tasks such as language execution, representation creation, coding, and information storage.

To elaborate further, typically, these models are trained on large-scale text corpora compiled from various online sources. During pre-training, these models learn linguistic structures through latent language modeling or self-supervised objectives such as subsequent token prediction. Subsequently, fine-tuning on smaller, domain-specific datasets enables adaptation to specialized fields such as medicine and law. This two-stage training process has become a standard method in contemporary NLP research. The system's pre-training and language processing steps have contributed to the development of more useful and effective models.

Wittgenstein and AI Language Models

We briefly mentioned Wittgenstein's ideas in the text. The philosopher's writings can be divided into two periods. What is important for our discussion is his emphasis on the meaning and context of language. We will briefly mention his observations and then try to position them in relation to our topic.

After a certain point, Wittgenstein radically changes his perspective. He no longer argues that language is a perfect logical system, but rather that it is a practical tool used by people in everyday life. His famous slogan is: "Meaning is use". He opposes the idea that language must be a flawless scientific order, because language already works seamlessly in our daily lives. Puzzles and confusion begin when we detach language from its normal use and "take it on vacation". Wittgenstein uses the example of "games" to prove that concepts do not share a single common essence. You cannot find a single feature that encompasses all games such as chess, football, and blind man's buff; however, there are "family resemblances" between them, just like members of a family. This shows that language also does not have a single essence, but consists of many parts (language games) that serve different functions. It clearly demonstrates Wittgenstein's transition from rigid logic to a flexible and practical philosophy of "way of life". Although the idea that "the limits of my language are the limits of my world" is particularly compelling, it later evolved into the notion that "the use of language is the meaning of my world" (14,15,27).

To understand and interpret language-centered approaches in AI, it is necessary to extend Wittgenstein's philosophy of language a little further. We previously discussed the philosopher's early and later thoughts. In his early work, Tractatus Logico-Philosophicus, Wittgenstein argued that language functions as a logical representation of facts in the world. According to this view, propositions derive their meaning from their correspondence with states of affairs, and understanding is based on logical structure (27-32). This perspective assumes that knowledge can be

represented through formal symbols and explicit rules, i.e., logical predicates and structured representation systems. Such an approach may offer a more explanatory framework for earlier symbolic AI systems based on formalism. However, it has been shown that such representations are insufficient for modeling the ambiguity, contextual variability, and pragmatic dimensions of natural language (15). Consequently, since language itself is a dynamic, constantly evolving, and context-dependent phenomenon, representation and symbolization alone may not be sufficient to fully grasp the functioning of AI models related to natural language. In general, the philosophy of language should offer a deeper and more nuanced perspective for understanding AI models. Indeed, the Philosopher has revised his views in later periods.

The philosopher abandoned the idea that language is shaped by a single fundamental logical structure in his later work, Philosophical Investigations. Instead, he argued that meaning arises from usage and emphasized the role of context, application, and social interaction. According to this perspective, language functioned as a process of "language games" in which words acquired meaning through their roles in specific activities and ways of life (14). This framework is further clarified by the concept of "family resemblance". In this approach, categories are defined not by fixed essences but by overlapping similarities between specific instances. In other words, meaning and context profoundly shape the dynamic nature of language. While words may appear to carry simple or isolated meanings on their own, within context they constantly transform into lived and situational realities. This approach is a more suitable explanation, particularly for ideal LLMs. It also paves the way for adapting human language abilities to AI models.

Now let's examine some of Wittgenstein's observations in the context of LLMs. Modern LLMs currently in use exhibit clear parallels with the framework expressed in Wittgenstein's later philosophy. Rather than relying on explicit symbolic definitions, LLMs derive meaning from usage patterns within broad linguistic contexts. Their capacity to generate contextually appropriate responses reflects a usage-based representation of language. However, these systems fall somewhat short of fully embodying Wittgenstein's concept of "forms of life". While LLMs can approximate linguistic behaviors, they still lack lived experience, cultural participation, and embodied understanding. Consequently, their linguistic competence remains disconnected from the social and experiential foundations that underpin the human process of meaning creation. These limitations become particularly apparent in fields such as psychotherapy, where context and meaning are deeply intertwined and central to practice. This is because fields like psychotherapy are heavily dependent on both context and the semantic reflections of the world.

Deep Learning, AI, Linguistics, and Cognitive Science

After questioning concepts such as meaning, representation, and context, it becomes possible to turn to the question of human cognition. Ultimately, it is the cognitive capacity of the human mind that produces and sustains these phenomena. Modern Homo sapiens have tremendous capacities related to language, learning, memory, and problem-solving. Perhaps AI will not be able to completely mimic these, but the systems that need to mimic them will be built on these very foundations. The natural course of development may also evolve the direction of progress in different ways. It should be reiterated that examining the relationship between AI architectures and human cognitive mechanisms can contribute both to improving system design and to theoretical insights into the nature of the human mind. However, pursuing this discussion in sufficient depth would exceed the scope of this article.

As previously mentioned, various mechanisms and processes found in deep learning models are crucial in the design of new AI models. Multi-task learning enables artificial systems to acquire multiple competencies simultaneously and transfer knowledge between tasks. Similarly, the human brain can combine visual cues during communication, interpret gestures, and adjust linguistic output according to social context. These abilities are crucial for domains based on human communication. Some researchers suggest that the brain's capacity for parallel and integrative information processing across cognitive domains exemplifies multi-task learning in artificial systems (32). In-depth examination and modeling of the human brain's cognitive production processes can significantly contribute to the further development of AI tools. This relationship is multifaceted: While human cognition influences the design and development of artificial systems, AI can also broaden our understanding of the depth and scope of human cognitive capacity. Furthermore, AI models can make significant contributions to neurophysiological research on the brain.

Applying pre-training and fine-tuning processes to AI models is also quite important. These processes reveal another significant similarity between artificial and human

learning processes. While models acquire general linguistic and conceptual knowledge through pre-training on large datasets, fine-tuning enables specialization in specific fields or tasks. This process is similar to how modern humans learn and become specialized in certain fields. Individuals also acquire broad foundational knowledge before specializing in specific areas such as medicine or legal linguistics (33). Initially, knowledge is encoded, contextualized, and learned through representational structures; this acquired knowledge is then applied in various situations. Just as lawyers conclude cases from the knowledge they have previously internalized, LLM systems can also perform domain-related inference operations thanks to this training and architecture.

Humans possess the capacity to learn from extensive sources of information, as well as the ability to acquire new concepts from relatively limited examples, and generally generate solutions to new problems through reasoning rather than relying on extensively stored knowledge. Artificial systems also have the capacity to make broad inferences, at least partially, with a small number of examples and through learning techniques. This process is similar to the working principles of language. Similarly, LLMs have demonstrated the ability to produce appropriate responses to tasks with minimal or no direct training data (34). Despite these advances, semantic fragility remains a significant limitation: LLMs can produce contextually inappropriate or logically inconsistent responses. As previously mentioned, human linguistic understanding relies not only on statistical regularities but also on contextual awareness, intent, cultural knowledge, and concrete experience. At this point, this fundamental limitation becomes evident once again. Further developments may help overcome this problem.

Neurophysiologically, the human mental process is a multifunctional mechanism that brings together phenomena into holistic representations. Meaning emerges through the integration of multiple modalities, including linguistic, visual, auditory, and tactile information. For example, the concept of "cat" is associated not only with a lexical label but also with sensory experiences, behavioral patterns, and lived interactions. Our cat at home is no longer just an ordinary cat. While the concept of a cat requires a certain processing, our cat at home is formed by the involvement of different mental functions in the process. Current AI systems exhibit only a limited capacity for this kind of multimodal integration. However, some artificial neural networks have been shown to develop internal structures resembling grammatical patterns during unsupervised learning and to reflect certain aspects of early human language acquisition (35). Despite these developments, it can be said that this artificial capacity is inferior to the productions of the human mind.

Some studies also continue to argue that human-like language capacity cannot be achieved solely through the use of large amounts of data (36). As highlighted in other content of this article, meaning in human cognition arises from the interaction of intelligent reasoning, contextual interpretation, and experience-based communication. By considering these dimensions together, it can be argued that linguistic theories can play a critical role in enhancing the meaningful capacities of AI outcomes. Including syntactic, meaningful, and pragmatic linguistic reports in the model architectures obtained from AI can help overcome some of the limitations found in purely learning methods.

In general, integrating deep learning into the fields of linguistics and systematics can open up opportunities for efficient and transformative research. Hybrid strategic architectures, multimodal processing, and intelligent storage options embedded in linguistics can make AI models more efficient. At the same time, it enables AI sites to offer more flexible, interpretable, and human-like solutions. In this way, the long-standing goal of creating "machines that can think like humans" can become increasingly achievable.

The Concept of Cognitive Enhancement and AI from a Cognitive Science Perspective

The issue we will look at is how AI models can help improve human thinking skills. This includes things like memory, attention, learning, and making decisions (37). In this case, AI is a major tool for improvement. As mentioned before, some AI learning methods are similar to how humans think. For example, multitasking learning lets artificial systems process many pieces of information at the same time, which is like how the human brain can handle different tasks together (32). Similarly, the pre-training and fine-tuning method is like how people first learn general knowledge before focusing on specific skills (33). After looking at these points, we can ask: can certain AI models also help improve human thinking abilities?

When we ask this question, we realize that even though there are similarities, these methods are only partial copies of human thought, not complete models. This comparison has its challenges. It's possible to improve certain thinking skills and develop certain parts of our cognitive ability using AI tools. However, right now, these capabilities are

limited. We can't predict what will happen in the future. Interestingly, some artificial neural networks have shown they can learn language structures similar to grammatical patterns through unsupervised learning. This is like how young children learn language (35). This ability has potential for more development. Using these advances in LLMs models could lead to better results. We could make more sophisticated neural networks.

One of the most effective and widely talked about uses of AI in improving thinking is the human-loop approach. In this model, AI systems support rather than take over thinking. They combine the speed of computers with human judgment and understanding of context. Supporting existing abilities is a more meaningful and lasting goal than replacing them. For example, in clinical decision support systems, AI can find patterns in big datasets while doctors handle the interpretation, ethical evaluation, and final decisions. From a cognitive science point of view, this working together model sees thinking not just as something that happens in the individual mind, but as a process that comes from interactions between people, technology, and the environment.

Transparency and understanding are important for AI to be a useful tool for improving thinking. Explainable AI (XAI) helps build trust, makes things more accountable, and supports better decisions by helping users understand how and why a system makes certain results (38). Without transparency, AI systems might become confusing black boxes, which can harm user trust and make people less willing to accept them ethically. Ethical issues around thinking improvement include becoming too reliant on automated systems, making existing biases worse, and possibly lowering human freedom (39). Handling these risks requires cooperation between technical experts, cognitive scientists, ethicists, and social scientists. When these conditions are met, AI models can greatly help human thinking abilities. This finding fits with the theory of distributed cognition (40,41). Cognitive science brings together knowledge from psychology, linguistics, neuroscience, and AI to model and understand thinking (42).

While some AI methods show aspects of human learning, there are still big gaps in areas like understanding context, emotional intelligence, and cultural meaning, often connected to the idea of "lifestyle" (36). A promising way to fill these gaps is through multimodal processing, which lets systems work with text, images, and sounds. These approaches are more like how humans make sense of meaning and can improve the understanding abilities of AI thinking tools (43). This deeper level of understanding could bring us closer to machines that can think, interpret, and maybe even experience things in ways that are similar to human thinking.

The Use of AI in the Psychiatric Field

Having addressed these conceptual areas, it is appropriate to briefly consider specific application areas. Among the most important of these are psychiatric and psychotherapeutic applications. After examining AI from the perspectives of representation, language, meaning, context, and cognitive enhancement, it becomes necessary to explore its concrete applications within psychiatric practice. Contemporary AI systems used in psychiatry are generally classified as Narrow AI, as opposed to hypothetical forms of Artificial General Intelligence or Super AI (9). These applications are designed to perform specific tasks in well-defined clinical contexts, such as symptom monitoring, risk estimation, or structured therapeutic support. Since our topic is not to discuss more advanced AI models that will be further developed in the future, we will only introduce the subject through existing narrow AI models.

From a methodological standpoint, some authors group AI systems used in psychiatry as symbolic, connectionist, and hybrid approaches. Symbolic systems include rule-based models that operate through pattern matching rather than actual semantic understanding. These include formal logic frameworks, expert systems, and early speech agents such as ELISA. Connectionist approaches encompass deep learning models used for speech analysis, image classification, and text-based prediction, often trained on large clinical or behavioral datasets. Hybrid models integrate symbolic reasoning with neural learning and are particularly valuable in clinical contexts where interpretability and transparency are essential. For example, they can combine rule-based constraints with neural classifiers to support clinical decision-making (3).

AI applications in psychiatry encompass a wide range of fields. Some researchers describe these as NLP-based models, those intertwined with data from assistive devices and digital platforms, often involving passive data collection and longitudinal analysis of language patterns, activity levels, and sleep rhythms to detect depressive symptoms, predict mood swings, and track treatment responses over time (44). Researchers have demonstrated that speech agents such as Woebot and Tess are effective in delivering structured cognitive-behavioral therapy

interventions. Randomized controlled trials and clinical evaluations have reported short-term reductions in symptoms of depression and anxiety among young adult users who interact with these tools through text-based dialogue (45). Some practitioners have also used machine learning techniques to identify early indicators of psychosis and schizophrenia. These include studies using automated speech analysis to predict the onset of psychosis in high-risk youth, as well as neuroimaging-based classifiers designed to differentiate clinical groups (46,47). In addition, the same researchers found that avatar-based therapies were effective in improving patients' ability to cope with auditory hallucinations. Randomized clinical trials show that repeated interaction with personalized virtual avatars can reduce the severity of paranoia and increase insight into hallucinatory voices (48). Taken together, these findings and developments appear truly promising.

Beyond diagnostic and therapeutic applications, some researchers have used social support robots like PARO to promote emotional engagement and social interaction in individuals with dementia. Cluster randomized controlled trials have reported improvements in agitation, mood, and social responsiveness during care sessions involving robotic interaction (49). Some clinicians have demonstrated similar effects in pilot studies examining applicability, user acceptance, and short-term mood changes during guided interactions. These clinicians have also shown that robotic applications like eBear are effective in older adults experiencing depressive symptoms (50). Continuous monitoring systems further contribute to psychiatric care by enabling early detection of high-risk behaviors and supporting longitudinal assessment through ongoing data collection.

The advantages of AI in psychiatry include scalability, cost-effectiveness, sustainable accessibility, and improved access to care, particularly in underserved or rural areas. It should also be noted that there are significant limitations in this field. These include a lack of genuine empathy, the emergence of algorithmic bias, risks related to data privacy, and unresolved issues concerning legal and professional accountability. Many authors highlight numerous of these problems in ethical discussions and critical analyses of AI use in healthcare (51,52). We should also point out that detailed and meaningful language use and cultural context are central realities in psychotherapy. Users must carefully design and utilize AI systems to avoid superficial or misleading interactions. Many of the issues we have discussed earlier are most strongly felt in the fields of psychiatry and psychotherapy. While advances in language modeling and contextual representation are expected, the field is still in its early stages of development.

In addition to clinical practice, AI has the potential to support psychiatric research, including large-scale randomized controlled trials. AI-based tools can assist in participant monitoring, data analysis, relapse prediction, and evaluation of treatment outcomes. Recent reviews highlight the role of AI in supporting, rather than transforming, clinical workflows by providing decision support and scalable interventions across a variety of psychiatric conditions (53,54). However, their integration into psychiatric practice must be guided by ethical oversight, clinical judgment, and human-centered care principles. The machines that "think and act like humans", envisioned by many thinkers since Turing's question, and perhaps even much earlier, may one day fully transform into machines that "conduct psychotherapy processes".

Discussion

This article addresses numerous topics. This is because AI is a reality with many dimensions. Most people view AI solely as a technical field of production. However, these tools have much deeper philosophical and intellectual aspects. Narrow perspectives, especially in the context of representation and meaning issues, are now insufficient. This is because these phenomena are the most important reflections of human language. As demonstrated throughout this study, the success of AI systems is directly related not only to computational power or data size, but also to how information is structured, within which representation framework it is produced, and how it is conveyed to the user. Therefore, XAI is not only a technical requirement with a high level of functionality, but also a tool based on knowledge production and suited to human language ability (55). Especially in clinical decision support systems, it is the way the output is justified, rather than the output itself, that generates trust. This trust is also quite important from an ethical point of view.

Repetition is not appropriate, but it seems inevitable to bring the issue together with the problem of representation. The issue of representation occupies a central position here. As some authors have pointed out (56), interpretability means not only internal transparency but also that the decision is contextually understandable. In a clinical context, this requires clearly showing which data patterns the diagnosis recommendation is based on. At this point, the representation problem ceases to be merely a technical

modeling issue; it moves into a broader framework concerning the construction of meaning, verifiability, and accountability. New paradigm searches such as quantum computing (57), while having the potential to increase representation capacity, may also make epistemological problems more complex rather than eliminating them. Therefore, the issue is not just more computation, but more qualified meaning production. Because the human mind is not a purely computational structure. Meaning is a reality constructed by this structure through language. In this respect, it again becomes the subject of philosophical debates.

The evolution of language models concretizes this debate. With the Transformer architecture, the capacity for capturing contextual patterns has increased significantly (23). However, this development also evokes a theoretical shift from Wittgenstein's early logical design understanding to his later "language games" approach (14). Modern LLMs can generate meaning from usage patterns to a limited extent; however, they still do not possess a "way of life". This is because they cannot fully represent the mind of a living human being. Above all, the question "Can machines really produce emotions?" remains unanswered. Who knows, perhaps one day AI models capable of feeling will be developed. For example, in fields such as psychiatry, meaning arises not only from word sequences but also from experience, culture, emotion, and contextual interaction. Therefore, the gap between technical progress and phenomenological experience persists. This situation seems likely to continue for quite some time.

The role of neuro-symbolic approaches in bridging this gap cannot be denied (30). Instead of purely statistical learning, the integration of logical inference and explicit representation layers can increase reliability, particularly in high-risk decision-making areas. The involvement of symbolic structures can contribute to making model outputs more consistent with clinical reasoning. However, this raises a new problem: Excessive formalization and algorithmization can diminish the flexibility of human communication. Therefore, the fundamental question for future research is how to establish a balanced structure between performance and depth of meaning.

Comparing linguistic models with neural networks is also important in this context. The differences between dependency parsers and neural disambiguation systems (58) demonstrate the extent to which artificial systems internalize the relationship between syntactic structure and semantic interpretation. The integration of structured lexical resources such as VerbNet (59) can increase contextual depth; however, the cultural layering of human language is still not fully represented. As mentioned earlier, all tools are still far from ideal. At this point, multimodality offers a new perspective. EEG-based language processing findings (60) and studies focusing particularly on the N400 component (61) demonstrate how the human brain processes semantic incongruity. Barsalou's embedded cognition approach (62) proposes that meaning is not solely linguistic but is also connected to bodily and action-based experience. Representations developed based on this approach could yield more reliable results, particularly in psychiatric applications. This is because psychiatric assessment requires not only verbal expression analysis but also the interpretation of tone, gestures, timing, and emotional context. The interpretation of bodily signals, in particular, is essential for social beings. Reading facial expressions, such as signs that appear when lying, is also important. Perhaps multimodal- neurobiological approaches can make significant contributions to the development of new AI models.

AI applications in psychiatry offer significant opportunities in response to growing demand (7,63). The need for accessibility has become more apparent in the post-pandemic period. However, it is clear that these tools should not replace human contact but rather support it. The real potential for transformation lies in the "cognitive partnership" model rather than "replacement". AI can support clinicians with its pattern recognition and data analysis capabilities; however, interpretation, ethical evaluation, and responsibility must remain the domain of human judgment.

The new perspective emerging from this study is as follows: The relationship between AI and fields such as psychiatry is not merely a matter of technological development; this relationship concerns how representation theories translate into clinical meaning production. In other words, the real issue is not "how accurate are AI predictions?" but rather "how well do the representations it produces align with clinical meaning?" Future research must evaluate the phenomenological adequacy of representations as well as performance metrics.

In conclusion, AI shows promise in fields such as psychiatric applications and psychotherapy; however, full utilization does not seem possible without resolving issues of representation, context, empathy, and explainability. Interdisciplinary collaboration, neuro-symbolic architectures, multimodal representations, and ethical

frameworks will be key determinants of this process. The claim to model the human mind requires not only technical progress but also a deeper questioning of the nature of meaning.

Footnotes

Authorship Contributions

Surgical and Medical Practices: A.R.T., Concept: F.B., A.R.T., Design: F.B., A.R.T., Data Collection or Processing: F.B., A.R.T., Analysis or Interpretation: F.B., A.R.T., Literature Search: F.B., A.R.T., Writing: F.B., A.R.T.

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

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

References

1. Turing AM. Computing machinery and intelligence. Mind. 1950;59(236):433-460.
2. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al. Language models are few-shot learners. Adv Neural Inf Process Syst. 2020;33:1877-1901.
3. Bender EM, Koller A. Climbing towards NLU: on meaning, form, and understanding in the age of data. In: Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics; 2020:5185-5198.
4. Devlin J, Chang MW, Lee K, Toutanova K. BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of NAACL-HLT. 2019;2019:4171-4186.
5. Lake BM, Baroni M. Generalization without systematicity: on the compositional skills of sequence-to-sequence recurrent networks. In: Proceedings of the 35th International Conference on Machine Learning (ICML 2018). 2018:2879-2888.
6. Schneider S, Leyer M. Me or information technology? Adoption of artificial intelligence in the delegation of personal strategic decisions. Manag Decis Econ. 2019;40:223-231.
7. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
8. World Health Organization. Mental Health Atlas 2020 [Internet]. 2021 [cited 2025 Aug 19]. Available from: https://www.who.int/publications/i/item/9789240036703
9. Russell SJ, Norvig P. Artificial intelligence: a modern approach. 4th ed. Hoboken, NJ: Pearson; 2021. Available link: http://lib.ysu.am/disciplines_bk/efdd4d1d4c2087fe1cbe03d9ced67f34.pdf
10. Nilsson NJ. Artificial intelligence: a new synthesis. San Francisco, CA: Morgan Kaufmann; 1998.
11. Simon HA, Newell A. Human problem solving. Englewood Cliffs, NJ: Prentice-Hall; 1971. Available link: https://iiif.library.cmu.edu/file/Newell_box00049_fld04211_doc0001/Newell_box00049_fld04211_doc0001.pdf
12. Shieber SM. The turing test: verbal behavior as the Hallmark of intelligence. Cambridge, MA: MIT Press; 2005;31(3).
13. Brachman RJ, Levesque HJ. Knowledge representation and reasoning. San Francisco, CA: Morgan Kaufmann; 2004. Available link: https://bista.sites.dmi.unipg.it/didattica/CSP/knowledge-representation-and-reasoning-the-morgan-kaufmann-series-in-artificial-intelligence.pdf
14. Wittgenstein L. Philosophical investigations. Oxford, UK: Basil Blackwell; 1953. Available link: https://static1.squarespace.com/static/54889e73e4b0a2c1f9891289/t/564b61a4e4b04eca59c4d232/1447780772744/Ludwig.Wittgenstein.-.Philosophical.Investigations.pdf
15. McCarthy J. Circumscription—a form of non-monotonic reasoning. Artif Intell. 1980;13(1-2):27-39.
16. Davis R, Shrobe H, Szolovits P. What is a knowledge representation? AI Mag. 1993;14(1):17-33.
17. Hebb DO. The organization of behavior: a neuropsychological theory. New York, NY: John Wiley & Sons; 1949. Available link: https://pure.mpg.de/rest/items/item_2346268_3/component/file_2346267/content
18. Rosenblatt F. The perceptron: a probabilistic model for information storage and organization in the brain. Psychol Rev. 1958;65(6):386-408.
19. Rumelhart DE, McClelland JL, eds. Parallel distributed processing: explorations in the microstructure of cognition. Vol 1. Cambridge, MA: MIT Press; 1986. Available link: https://gwern.net/doc/ai/nn/1986-rumelhart-pdp-v1.pdf
20. Goodfellow I, Bengio Y, Courville A. Deep learning. Cambridge, MA: MIT Press; 2016. Available link: https://repo.darmajaya.ac.id/4189/1/Deep%20learning_%20adaptive%20computation%20and%20machine%20learning%20%28%20PDFDrive%20%29.pdf
21. Newell A. Unified theories of cognition. Cambridge, MA: Harvard University Press; 1990. Available link: https://iiif.library.cmu.edu/file/Newell_box00022_fld01532_doc0001/Newell_box00022_fld01532_doc0001.pdf
22. Anderson JR. ACT: a simple theory of complex cognition. Am Psychol. 1996;51(4):355-365.
23. Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, et al. Attention is all you need. Adv Neural Inf Process Syst. 2017;30:5998-6008.
24. Jurafsky D, Martin JH. Speech and language processing [Internet]. 3rd ed. draft. 2025 [cited 2025 Aug 19]. Available from: https://web.stanford.edu/~jurafsky/slp3/
25. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735-1780.
26. Radford A, Wu J, Child R, Luan D, Amodei D, Sutskever I. Language models are unsupervised multitask learners. OpenAI Blog [Internet]. 2019 [cited 2025 Aug 19]. Available from: https://cdn.openai.com/better-language- models/language_models_are_unsupervised_multitask_learners.pdf
27. Wittgenstein L. Tractatus logico-philosophicus. Pears DF, McGuinness BF (trans). London, UK: Routledge; 2001. Originally published 1921.
28. Montague R. The proper treatment of quantification in ordinary English. In: Hintikka J, Moravcsik J, Suppes P (eds). Approaches to Natural Language. Dordrecht, NL: Reidel; 1973:221-242.
29. Frege G. Über sinn und bedeutung. Z Philos Philos Krit. 1892;100(1):25-50.

30. Besold TR, Garcez AA, Bader S, Bowman H, Domingos P, Hitzler P, et al. Neural- symbolic learning and reasoning: a survey and interpretation. arXiv [Internet]. 2017 [cited 2025 Aug 18].

31. Zambak A, Vergauwen R. Artificial intelligence and agentive cognition: a logico- linguistic approach. Logique Analyse. 2009;52(205):57-96.

32. Barrett DG, Morcos AS, Macke JH. Analyzing biological and artificial neural networks: challenges with opportunities for synergy? Curr Opin Neurobiol. 2019;55:55-64.

33. Flesch T, Balaguer J, Dekker R, Nili H, Summerfield C. Comparing continual task learning in minds and machines. Proc Natl Acad Sci U S A. 2018;115(44):E10313-E10322.

34. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014;15:1929-1958.

35. Hewitt J, Manning CD. A structural probe for finding syntax in word representations. In: Proceedings of NAACL-HLT 2019; 2019 Jun; Minneapolis, MN. Stroudsburg, PA: Association for Computational Linguistics; 2019:4129-4138.

36. Ravichander A, Belinkov Y, Hovy E. Probing the probing paradigm: does probing accuracy entail task relevance? arXiv [Internet]. 2020 [cited 2025 Aug 19].

37. Kaplan A, Haenlein M. Siri, Siri, in my hand: who's the fairest in the land? Bus Horiz. 2019;62(1):15-25.

38. Gunning D, Aha D. DARPA's explainable artificial intelligence (XAI) program. AI Mag. 2019;40(2):44-58.

39. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1(9):389-399.

40. Dellermann D, Ebel P, Söllner M, Leimeister JM. Hybrid intelligence. Bus Inf Syst Eng. 2019;61(5):637-643.

41. Hutchins E. Cognition in the wild. Cambridge, MA: MIT Press; 1995.

42. Miller GA. The cognitive revolution: a historical perspective. Trends Cogn Sci. 2003;7(3):141-144.

43. Kiela D, Bulat L, Vero AL, Clark S. Virtual embodiment: a scalable long-term strategy for artificial intelligence research. arXiv [Internet]. 2016 [cited 2025 Aug 19].

44. Inkster B, Sarda S, Subramanian V. An empathy-driven conversational AI agent (Wysa) for digital mental well-being. JMIR Mhealth Uhealth. 2018;6(11):e12106.

45. Fitzpatrick KK, Darcy A, Vierhile M. Delivering CBT using a fully automated conversational agent (Woebot). JMIR Ment Health. 2017;4(2):e19.

46. Vieira S, Pinaya WHL, Mechelli A. Using deep learning to investigate neuroimaging correlates of psychiatric disorders. Neurosci Biobehav Rev. 2017;74(Pt A):58-75.

47. Bedi G, Carrillo F, Cecchi GA, Slezak DF, Sigman M, de la Iglesia H, et al. Automated analysis of free speech predicts psychosis onset. NPJ Schizophr. 2015;1:15030.

48. Garety P, Ward T, Emsley R, Greenwood K, Freeman D, Fowler D, et al. Effects of slowMo on paranoia. JAMA Psychiatry. 2021;78(7):714-725.

49. Moyle W, Jones CJ, Murfield JE, Thalib L, Beattie ERA, Shum DKH, et al. Robotic seal therapy in dementia. J Am Med Dir Assoc. 2017;18(9):766-773.

50. Kargar BAH, Mahoor MH. eBear socially assistive robot pilot study. In: 2017 IEEE- RAS Int Conf Humanoid Robots; 2017.

51. Floridi L, Cowls J, Beltrametti M, Chatila R, Chazerand P, Dignum V, et al. AI4People ethical framework. Minds Mach. 2018;28(4):689-707.

52. Hariman K, Patel J, Whiting P. Artificial intelligence in psychiatry: are we ready yet? BJPsych Adv. 2019;25(5):279-286.

53. Pham KT, Nabizadeh A, Selek S. Artificial intelligence and chatbots in psychiatry. Psychiatr Q. 2022;93(1):249-253.

54. Ray A, Bhardwaj A, Malik YK, Singh S, Gupta R. Artificial intelligence and psychiatry: an overview. Asian J Psychiatr. 2022;70:103021.

55. Gunning D. Explainable artificial intelligence (XAI). DARPA [Internet]. 2017 [cited 2025 Aug 18]. Available from: https://www.darpa.mil/program/explainable-artificial- intelligence

56. Doshi-Velez F, Kim B. Towards a rigorous science of interpretable machine learning. arXiv [Internet]. 2017 [cited 2025 Aug 18]. Available from: https://arxiv.org/abs/1702.08608

57. Biamonte J, Wittek P, Pancotti N, Rebentrost P, Wiebe N, Lloyd S. Quantum machine learning. Nature. 2017;549(7671):195-202.

58. Kiperwasser E, Goldberg Y. Dependency parsing using bidirectional LSTM features. arXiv [Internet]. 2016 [cited 2025 Aug 19].

59. Schuler KK. VerbNet: A broad-coverage, comprehensive verb lexicon [PhD dissertation]. Philadelphia (PA): University of Pennsylvania; 2005.

60. Luck SJ. An introduction to the event-related potential technique. 2nd ed. Cambridge, MA: MIT Press; 2014.

61. Kutas M, Federmeier KD. Thirty years of the N400. Annu Rev Psychol. 2011;62:621-647.

62. Barsalou LW. Grounded cognition. Annu Rev Psychol. 2008;59:617-645.

63. Belli H, Laçin S. Artificial intelligence in psychiatry: applications and challenges. Atlas J Med. 2025;5(14):224-233.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):177-184
DOI: 10.4274/BMB.galenos.2025.87597

The Psychosocial Status of Physically and/or Intellectual Disability Children's Siblings

Fiziksel ve/veya Zihinsel Engelli Çocukların Kardeşlerinin Psikososyal Durumu

Demet Yılmaz[1], Dilek Toprak[2], Gülşen Köse[3]

[1]İstanbul Şişli Family Healthcare Center, Clinic of Family Medicine, İstanbul, Turkey
[2]İstanbul Atlas University Faculty of Medicine, Department of Family Medicine, İstanbul, Turkey
[3]İstinye University Live Hospital, Department of Pediatric Neurology, İstanbul, Turkey

Abstract

Objective: In this study, we aimed to evaluate psychosocial status, the relationship between socio-demographic factors, to support the psychosocial development of physically and/or mentally disabled children's siblings that can be often overlooked, we also aim to plan new projects, researches and education that will support families and children in the light of our findings.

Method: This study was planned as a one centered, a cross-sectional study with prospectively recruited participants. The children aged 6-17 years who had physical and/or mental retarded siblings, admitted to University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Resarch Hospital, Family Medicine Clinic and Pediatric Neurology Clinic were included in the study. Children's psychosocial status and depression levels were questioned face to face interview with children-adolescent's depression inventory (CDI) and the results were analyzed using frequency, chi-square and t-test.

Results: Most of the participants of 51 children in the study, were girl (60.8%) and mean age was 11.25±3.14. The average CDI score was 11.76±6.56 for the healthy siblings. 15.7% of children were depressed and 75% of them were girls. Average CDI scores of children with depression was found as 22.87±3.90. Mean CDI scores of girls (12.03±6.32) compared with boys (11.35±7.06) was higher, however, this relationship was not statistically significantly. The average CDI scores (13.5±7.21) of subjects whose siblings do not get education were higher than those who get education (11.23±6.35). While 12.8% of children whose siblings got education, had depression; 25.0% of siblings who have not educated were depressed (p>0.05).

Conclusion: This study showed that siblings of disabled children may experience psychosocial problems and are at risk of developing

Öz

Amaç: Bu çalışmada, çoğu kez göz ardı edilebilen fiziksel ve/veya zihinsel engelli çocukların kardeşlerinin psikososyal durumlarını ve sosyo-demografik özellikleriyle ilişkisini değerlendirmeyi, çocukların psikososyal gelişimlerinin desteklenmesini; elde ettiğimiz bulgular ışığında aileye ve çocuklarına destek olacak yeni projeler oluşturulmasını amaçladık.

Yöntem: Çalışma katılımcıların prospektif olarak alındığı kesitsel, tek merkezli, tanımlayıcı ve analitik nitelikte bir çalışma olarak planlandı. Sağlık Bilimleri Üniversitesi, Şişli Hamidiye Etfal Eğitim ve Araştırma Hastanesi, Aile Hekimliği Polikliniği ve Çocuk Nörolojisi Polikliniği'ne başvuran, fiziksel ve/veya mental retarde çocukları olan ailelerin 6-17 yaş arası sağlıklı çocukları dahil edildi. Çocukların psikososyal durumu ve depresyon düzeyleri çocuk-ergen depresyon ölçeği (ÇDÖ) uygulanarak, yüzyüze görüşmeyle değerlendirildi. Analizlerde frekans, ki-kare, t-testi kullanıldı; p<0,05 anlamlı kabul edildi.

Bulgular: Çalışmaya katılan 51 çocuğun çoğu kız (%60,8) ve yaş ortalaması 11,25±3,14 idi. Sağlıklı kardeşlerin ortalama ÇDÖ puanı 11,76±6,56 idi. %15,7'sinin depresyonda olduğu; depresyonda olanların %75'inin kız çocuk olduğu belirlendi. Depresyonda olan çocukların ortalama ÇDÖ puanı 22,87±3,90 olarak bulundu. Kız çocukların ortalama ÇDÖ puanının (12,03±6,32), erkek çocuklara göre (11,35±7,06) daha yüksek olduğu ancak aralarında anlamlı bir ilişki olmadığı saptandı. Gruplar arasında anlamlı bir ilişki saptanmasa da, kardeşi eğitim almayan çocukların ortalama ÇDÖ puanı (13,5±7,21), kardeşi eğitim alanlara göre (11,23±6,35) daha yüksekti. Bununla birlikte kardeşi eğitim alanların %12,8'inin, kardeşi eğitim almayanların ise %25,0'ının depresyonda olduğu saptandı (p>0,05).

Address for Correspondence: Demet Yılmaz, İstanbul Şişli Family Healthcare Center, Clinic of Family Medicine, İstanbul, Turkey
E-mail: 24demetyilmaz@gmail.com **ORCID:** orcid.org/0000-0001-6119-1734
Received: 04.08.2024 **Accepted:** 26.11.2025 **Epub:** 04.12.2025 **Publication Date:** 24.06.2026

Cite this article as: Yılmaz D, Toprak D, Köse G. The psychosocial status of physically and/or intellectual disability children's siblings. Bagcilar Med Bull. 2026;11(2):177-184

Abstract

depression. The developmental period of children with disabled siblings is of great importance and should not be overlooked.

Keywords: Children, depression, disabled, psychosocial development, sibling

Öz

Sonuç: Bu çalışma, engelli çocukların kardeşlerinin psikososyal gelişimleriyle ilgili problemler yaşayabileceklerini, depresyona meyilli ya da depresyonda olabileceklerini göstermiştir. Engelli çocukların sağlıklı kardeşlerinin gelişim süreci göz ardı edilemeyecek kadar büyük önem taşımaktadır.

Anahtar kelimeler: Çocuklar, depresyon, engelli, kardeş, psikososyal gelişim

Introduction

The needs of siblings often compete with those of their ill brothers or sisters, which increases the risk that they may suppress or neglect their own needs (1). Research indicates that siblings of children with chronic illnesses experience both positive and negative consequences related to their family situation (2). Negative outcomes include increased emotional and behavioral difficulties, as well as feelings of fear, anxiety, sadness, anger, guilt, helplessness, and uncertainty. Additional challenges may involve family disruptions, limited parental support, school difficulties, changes in friendships, unmet needs, and a reduced health-related quality of life (2-5).

Erik H. Erikson proposed a theory of psychosocial development (6) based on the "principle of gradual formation (epigenetics)" (7,8). According to Erikson, personality development occurs in eight overlapping stages in which conflicts between positive and negative elements (such as basic trust and mistrust) emerges at each stage evolve and mature. The tension between these opposing forces defines the stage-specific crisis and ultimately determines which emotion will predominate (9).

Erikson's eight psychosocial developmental stages are as follows:

1. Basic trust versus distrust (0-1 years)
2. Autonomy versus doubt and shame (2-3 years)
3. Initiative versus guilt (3-5 years)
4. Industry (success) versus inferiority (6-11 years)
5. Identity versus role confusion (11-20 years)
6. Intimacy versus isolation (young adulthood)
7. Generativity versus stagnation (adulthood)
8. Ego integrity versus despair (old age) (10).

Depression in Children

The term "depression" refers to a persistent state of low mood. Depressive disorders in children and adolescents represent a significant public health concern because of their high prevalence and persistent negative effects on cognitive, social, and psychological development. These disorders affect approximately 2-3% of children and 8% of adolescents (11).

Symptoms and Diagnostic Criteria of Depression in Children

Depression in children can be difficult to detect because its onset is often vague. Depressive symptoms in children can present differently and present with different clinical manifestations. In early childhood, symptoms are particularly variable and are often misinterpreted as behavioral issues. These may include hyperactivity, irritability, self-harm or environmental damage, introversion, emotional instability, and sudden mood swings (12).

Among school-aged children, depression may present as persistent sadness or irritability, diminished interest in previously enjoyable activities, changes in appetite (resulting in weight fluctuations), insomnia or hypersomnia, psychomotor agitation or retardation, fatigue, feelings of worthlessness or guilt, reduced attention span, and in severe cases, suicidal ideation (13).

According to the diagnostic and statistical manual of mental disorders-5, a diagnosis of major depressive disorder requires the presence of at least five symptoms lasting for a minimum of two weeks. One of these must be a depressed mood, an irritable mood, or a loss of interest or pleasure (14).

Additional symptoms may include:

1. Lack of expected weight gain
2. Insomnia or excessive sleep
3. Psychomotor agitation or retardation
4. Fatigue or loss of energy
5. Feelings of worthlessness or excessive guilt
6. Reduced ability to think or concentrate
7. Recurring thoughts of death.

These symptoms must cause clinically significant impairment in social or academic functioning (11).

This study aims to evaluate the psychosocial status of siblings of children with physical and/or mental disabilities—an often overlooked group—and to propose projects, research, and educational initiatives to support these families and their children.

Materials and Methods

This study was performed with the approval of the University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital Clinical Research Ethics Committee (date: 15/03/2016; approval number: 1125).

Characteristics of the Study

This study was planned as a single-center, descriptive and analytical cross-sectional study with prospectively recruited participants. The study population included patients who visited the family medicine and pediatric neurology outpatient clinics at the training and research hospital between March 30 and June 10, 2016. Sample size was calculated using the Yamane formula. Statistical calculations indicated that at least 40 children must be included in the study. A total of 51 healthy children (aged 6-17 years) who met the inclusion criteria and consented to participate, from families of children with physical and/or mental disabilities, were included (n=51).

Prior to participation, both the children and their parents were informed in detail about the study's objectives. Parents completed a socio-demographic questionnaire, while the child-adolescent depression inventory (CDI) was administered to children via face-to-face interviews. The obtained CDI scores were subsequently analyzed.

CDI and Evaluation

The children's depression scale was developed by Kovacs (15) on the basis of Beck's depression inventory, through the addition of depression-specific relations. The validity and reliability study for the Turkish population was carried out by Öy (16).

Children's psychosocial development and depression levels were assessed using the CDI, a self-assessment tool suitable for children aged 6-17 years. The CDI consists of 27 items written in simple language to ensure understanding by this age group. The inventory can be completed by the child independently or with assistance and offers three response options for each item.

For example:

1. "I feel sad from time to time"

2. "I often feel sad"

3. "I always feel sad".

Each item is scored based on symptom severity: 0, 1, or 2 points. Children are asked to reflect on their feelings over the past two weeks and select the most appropriate statement. The maximum possible score is 54 with a cut-off point of 19. Compared with 19, the higher the obtained score is, the more severe the depression can be considered (16).

Statistical Analysis

The data were analyzed using SPSS (Statistical Package for the Social Sciences) version 20.0. Descriptive statistics included mean, standard deviation, median, minimum, maximum, frequency, and ratios. The distributions of the variables were assessed using the Kolmogorov-Smirnov test. Kruskal-Wallis and Mann-Whitney U tests were used in the analysis of quantitative data. The chi-square test was used to analyze qualitative data, and Spearman correlation analysis was used to assess correlations. Significance was evaluated at the $p<0.05$ level.

Results

A total of 51 children who had siblings with physical and/or mental disabilities were included in the study; the majority were female (60.8%, n=31). The ages of the participants ranged from 6 to 17 years, with a mean age of 11.25±3.14 years. By age group, 26 children (51%) were aged 6-11 years and 25 (49%) were aged 11-17 years. Sixteen children (31.4%) had one sibling, 19 children (37.3%) had two siblings, and 16 children (31.4%) had three or more siblings. Among the disabled siblings, ages ranged from 1 to 24 years; the mean duration of illness was 7 years, and the mean age at treatment initiation was 1 year.

Most parents were married (94.1%; n=48). Mothers' educational levels were as follows: Primary school graduates, 56.9% (n=29); university graduates, 11.8% (n=6); literate, 3.9% (n=2); and illiterate 2% (n=1). Fathers' educational levels were as follows: Primary school graduates, 56.9% (n=29); university graduates, 11.8% (n=6); and illiterate, 2% (n=1). Similarly, 2% (n=1) of the fathers were illiterate. While 92.2% (n=47) of mothers were not working, 98.2% (n=50) of fathers were working. The average CDI scores of children, stratified by socio-demographic characteristics, are presented in Table 1. 7.8% (n=4) of the children had

a family history of psychiatric illness, and their rates of depression were significantly higher than those without a family history (Table 2).

Among the children, 47% (n=24) had difficulty sleeping, 41.2% (n=21) had a poor appetite, and 33.3% (n=17) often worried that something terrible would happen. Additionally, 37.3% (n=19) often worried about aches and pains, and 68.6% (n=35) believed that they were responsible for negative events. In terms of social interactions, 19.6% (n=10) did not like being around people most or all of the time; 41.2% (n=21) often felt lonely; and 47.0% (n=24) reported having few or no friends. School performance was another area of concern: 54.9% (n=28) showed poor or deteriorating academic performance.

Among siblings with disabilities, 9.8% (n=5) had physical disabilities only, 43.1% (n=22) had mental disabilities only, and 47.1% (n=24) had both physical and mental disabilities. The breakdown of disability types is shown in Graphic 1. When the educational status of the disabled siblings was evaluated, 76.5% (n=39) were receiving an education, while 23.5% (n=12) were not (Graphic 2). Although no statistically significant relationship was found between CDI scores and sibling education status (p=0.225), children with siblings not receiving an education had higher average CDI scores (13.5±7.21) than those with siblings receiving an education (11.23±6.35). Among children with disabled siblings who were educated, 12.8% were depressed, compared with 25.0% of children whose disabled siblings were not receiving an education (p=0.310), as shown in Table 3.

The overall mean CDI score among participants was 11.76±6.56, with 15.7% (n=8) meeting the criteria for depression. Among the depressed children, 75% (n=6) were girls. The average CDI score was 22.87±3.90 in depressed children and 9.69±4.56 in non-depressed children. Analysis by age showed that 6-year-olds had the highest mean CDI score (19.66±9.29), while 11-year-olds had the lowest (8.0±5.14). When examining scores by age group and gender, children aged 12 to 17 and girls had higher CDI scores than younger children and boys, although these differences were not statistically significant.

The mean score of children with physically disabled siblings was 10.20±6.83; the mean score of children with mentally disabled siblings was 11.77±6.09. Although children with

Table 1. Mean CDI scores according to socio-demographic factors

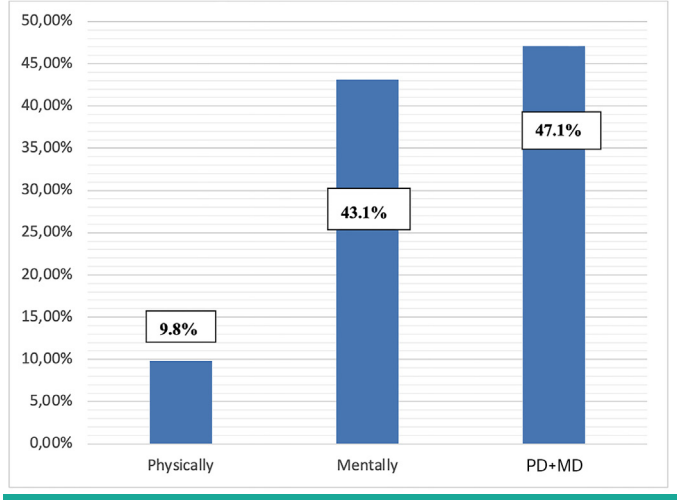
	n	Mean CDI score	p-value
Gender			0.463
Girl	31	12.03±6.32	
Boy	20	11.35±7.06	
Age			
6-11 years	26	10.92±6.64	0.294
12-17 years	25	12.64±6.49	
Number of siblings			
Has 1 sibling	16	10.43±6.21	
Has 2 siblings	19	12.31±6.78	0.223
Has 3 and more siblings	16	12.43±6.85	
Mother's education level			
Uneducated, primary or secondary school	34	12.14±6.65	0.371
High school or university	17	11.0±6.50	
Father's education level			
Uneducated, primary or secondary school	35	12.42±6.74	0.358
High school or university	16	10.31±6.10	
Mother's working status			
Works	4	16.25±1.89	0.221
Does not work	47	11.38±6.68	

Mann-Whitney U test, Kruskal-Wallis H test, CDI: Children-adolescent's depression inventory

Table 2. The relationship between the family history of psychiatric illness and children's depression status

		Depression status		p-value
		Yes	No	
Family history of psychiatric illness	Yes	50.0% (n=2)	50.0% (n=2)	0.049
	No	12.8% (n=6)	87.2% (n=41)	

chi-square test



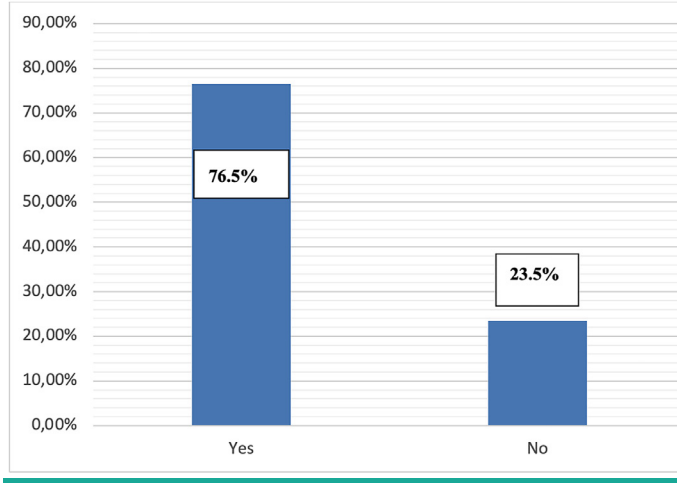
Graphic 1. Distribution of disability type of the disabled siblings

PD+MD: Physically and mentally disability

siblings who had both physical and mental disabilities recorded the highest mean scores (12.08±7.13), there was no statistically significant difference in CDI scores by sibling disability type.

When the children's answers were evaluated, 8.8% sometimes feared that something bad would happen and felt sad, and 4.7% sometimes wanted to cry and felt sad all the time. As shown in Table 4, a significant correlation was found between these groups (p=0.049; p=0.014).

All children who were self-loathing also felt ugly (p=0.002). Among children who reported being "always bored", 50.0% described themselves as "bad" children, as did 42.9% of those who often felt tired, 22.7% of those experiencing academic decline, and 22.2% of those frequently in conflict with others; the associations were statistically significant (p=0.018, 0.002, 0.026, and 0.044, respectively).



Graphic 2. Education status of disabled siblings

Discussion

Few studies have focused on siblings' well-being (17). This study aims to assess the psychosocial development and depressive symptoms of siblings of children with disabilities. A proportion of participants scored above the CDI cut-off, suggesting that they may be at an increased risk of depressive symptoms. Siblings of children with disabilities represent an at-risk group (18). Recent reviews and meta-analyses have shown that siblings have an increased risk of mental health problems and decreased well-being compared with their peers (18-20).

When CDI scores were analyzed by age, six-year-old children exhibited the highest mean scores. Younger children have heightened developmental needs for love, attention, play, and shared time with their parents. This finding suggests that depressive symptoms in healthy siblings may be more pronounced when these fundamental needs are not met, particularly during early childhood. Supporting this conclusion, studies by Powell and Ogle (21), Valdivieso et al. (22) Bayhan and Yükselen (23) show that younger siblings may have difficulty interpreting events realistically. They noted that non-disabled siblings may perceive the time and attention given to their disabled siblings as a form of parental rejection, which may lead to feelings of resentment and anger. In addition, they may wonder why their parents favor their disabled siblings and even imitate their siblings' physical or behavioral characteristics (21-23).

Our study found no significant difference in CDI scores between boys and girls, which is consistent with Senel's (24) findings. In Senel's (24) study titled "Comparison of the Attitudes towards Disability and Anxiety Levels

Table 3. The relationship between the educational status of the disabled siblings of children and the depression status of non-disabled children

		Depression status		p-value
		Yes	No	
Disabled sibling's education status	Having education	12.8% (n=5)	87.2% (n=34)	0.310
	Not having education	25.0% (n=3)	75.0% (n=9)	
chi-square test				

Table 4. The relationship between frequency of thinking something bad will happen and feeling sad of children

	Those who feel sad from time to time	Those who often feel sad	Those who always feel sad	p-value
Those who think that something bad will happen to them from time to time	88.2% (n=30)	8.8% (n=3)	2.9% (n=1)	0.049
Those who cry from time to time	79.1% (n=34)	16.3% (n=7)	4.7% (n=2)	0.014
chi-square test				

of Those Who Have and Do Not Have Siblings with Disabilities", which involved 30 healthy siblings and 30 siblings diagnosed with intellectual disabilities, autism, or Down syndrome, no significant differences in anxiety levels were found according to gender, family size, or parental education level.

Although there was no statistically significant correlation between CDI scores and the type of disability (physical, mental, or both), siblings of children with both physical and mental disabilities had the highest average CDI scores. Having a sibling with a disability increases the need for care and support, affects the educational, social, and personal needs of the healthy sibling, consequently affecting his/her psychosocial development.

Regarding the educational status of disabled siblings, children whose disabled siblings did not receive education had higher average CDI scores than children whose disabled siblings received education. Depression rates were approximately twice as high among children whose siblings with disabilities did not receive an education. Ertürk (25) found significant differences in "Sibling Problems Questionnaire'' scores between siblings of disabled and non-disabled children in a study of siblings in integrated education programs.

In our study, nearly half of the children had sleep disturbances and loss of appetite, and more than half experienced decreased academic success.

Similarly, Sloper (26) found that siblings of children with chronic illnesses often experienced significant eating and sleep disturbances following the child diagnosis. In their study of depressive symptoms among siblings of children with disabilities, Martinez et al. (18) reported that siblings of children with chronic health problems were at greater risk of symptoms suggestive of depression than siblings of healthy children.

Dinkelbach et al. (2) reported decreased physical and psychological well-being and lower self-esteem among siblings of children and adolescents with life-limiting conditions who were receiving pediatric palliative care at home. Siblings of disabled or chronically ill children are often overlooked members of the family, and their needs for support may go unmet (27). Siblings often have little knowledge about disabilities and may feel ostracised (28). In one study, many participants, who were healthy siblings of people with additional or complex needs, reported a disparity in attention received between themselves and their sibling. As we can see below, this often came from immediate or extended family members or other adults they knew, even during difficult times such as the death of a family member:

"Nobody really asked me anything." When my mum died, everyone just kept asking me how my sister was coping with my mum's death; they asked me how my dad was or how my sister was, but nobody actually asked me how I was (27). Additionally, Wawrzynski et al. (29) found that siblings often experience differential treatment by parents, friends, and others, and consequently feel more isolated and restricted in their interactions.

Most studies have reported that a sibling's disability negatively affects sibling functioning, causing increased anxiety and depression, social withdrawal, conduct disorders, and other behavioral problems (17).

In studies examining emotional and behavioral adjustment, the primary internalizing and externalizing symptoms identified were hyperactivity, hostility, anxiety, and somatic and emotional problems (30-36). In one study, more than 50% of siblings of individuals with disabilities reported feelings of guilt, described their sibling relationships as poor, exhibited elevated levels of depressive symptoms, and demonstrated reduced overall well-being (37).

It is crucial not to neglect the health needs of my siblings. Supporting the psychosocial development of these children is essential for their future well-being. Studies have shown that adult siblings of individuals with intellectual and developmental disabilities often report higher levels of depression and anxiety, poorer sibling relationships, and lower life satisfaction (38).

Study Limitations

In our study, the cases were evaluated only once. More data can be obtained from studies that include larger numbers of participants and that regularly evaluate healthy siblings with psychosocial support. It would be more useful to evaluate developmental processes comparatively in follow-up assessments during childhood, adolescence, and adulthood.

Conclusion

Our study indicates that siblings of children with disabilities may experience challenges in their psychosocial development and be prone to depressive symptoms or clinical depression. In particular, depressive symptoms were apparent in very young children, possibly because parents were unable to fully meet the healthy sibling's needs.

When educating parents on caring for children with disabilities, practitioners should also focus on the psychosocial well-being of healthy siblings by providing them with concrete, practical guidance. Considering that siblings' anxiety levels may decrease following support programs, it is recommended to provide psychological support to siblings of children with disabilities and to establish structured support programs.

The needs of siblings of children with disabilities remain under-researched in current medical practice, with limited scientific studies addressing this issue. More research and interventions are needed to support these families by recognizing siblings as essential family members and future contributors to society. In light of these findings, new projects, research, and educational programs should be developed to ensure that these children do not feel excluded.

Ethics

Ethics Committee Approval: This study was performed with the approval of the University of Health Sciences Turkey, Şişli Hamidiye Etfal Training and Research Hospital Clinical Research Ethics Committee (date: 15/03/2016; approval number: 1125).

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

Footnotes

The study was presented at 22nd WONCA Europe Conference Prague, Czech Republic, June 28-July, 2017 (poster presentation).

Authorship Contributions

Concept: D.Y., D.T., G.K., Design: D.Y., D.T., G.K., Data Collection or Processing: D.Y., D.T., G.K., Analysis or Interpretation: D.Y., D.T., G.K., Literature Search: D.Y., D.T., G.K., Writing: D.Y., D.T., G.K.

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

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

References

1. Kittelsen TB, Castor C, Lee A, Kvarme LG, Winger A. "What about me?": lived experiences of siblings living with a brother or sister with a life-threatening or life-limiting condition. Int J Qual Stud Health Well-being. 2024;19(1):2321645.
2. Dinkelbach L, Köhler M, Galushko M, Pieper L, Kuhlen M, Danneberg M, et al. Psychosocial well-being of siblings of pediatric patients in palliative home care. J Pain Symptom Manage. 2023;66(6):630-637.e1.
3. Fullerton JM, Totsika V, Hain R, Hastings RP. Siblings of children with life-limiting conditions: psychological adjustment and sibling relationships. Child Care Health Dev. 2017;43(3):393-400.
4. Alderfer MA, Long KA, Lown EA, Marsland AL, Ostrowski NL, Hock JM, et al. Psychosocial adjustment of siblings of children with cancer: a systematic review. Psychooncology. 2010;19(8):789-805.
5. Long KA, Lehmann V, Gerhardt CA, Carpenter AL, Marsland AL, Alderfer MA. Psychosocial functioning and risk factors among siblings of children with cancer: an updated systematic review. Psychooncology. 2018;27(6):1467-1479.
6. Arslan E. (Advisor: Arı R.). Investigation of Erikson's psychosocial development periods and ego identity processes in adolescents in terms of attachment styles. Doctoral Thesis, Institute of Social Sciences, Selcuk University. Department of child development and home management; child development and education department, Konya:2008.
7. Erik H. Erikson. Identity, youth and crisis. New York: W. W. Norton Company, 1968.
8. Erik H. Erikson. Identity and life cycle. New York: W. W. Norton Company, 1980.
9. Dereboy IF. Identity confusion, understanding, recognizing, handle. 1st ed., Malatya: Ozmert Ofset, 1993.
10. Erikson EH. The eight period of huma. Ankara: Individual and Community Publishing, 1984.
11. Kaplan BJ, Sadock VA, Ruiz P. Synopsis of psychiatry behavioral sciences/clinical psychiatry. 11st ed., Lippincot Williams & Wilkins. [Psikiyatri: davranış bilimleri/klinik psikiyatri. (Ali Bozkurt, Çev. Ed.). Ankara: Güneş Tıp Kitabevleri]; 2016.
12. Long C. (Advisor: Deniz ME.). Evaluation of depression and social skills in children with and without divorced parents. Master's Thesis, Istanbul Arel University Institute of Social Sciences, Department of Psychology, Istanbul:2013.
13. Taner YI. Depressive and bipolar disorders in children, adolescents, adults and the elderly. Department of Child and Adolescent. Ankara, 2013.
14. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. [Internet]. Available from:https://www.psychiatry.org/psychiatrists/practice/dsm.
15. Kovacs M. The children's depression, inventory (CDI). Psychopharmacol Bull. 1985;21(4):995-998.
16. Öy B. A validity and reliability study of the depression scale for children. Turkish Journal of Psychiatry. 1991;2(2):132-136.
17. Levante A, Martis C, Del Prete CM, Martino P, Primiceri P, Lecciso F. Siblings of persons with disabilities: a systematic integrative review of the empirical literature. Clin Child Fam Psychol Rev. 2025;28(1):209-253. Erratum in: Clin Child Fam Psychol Rev. 2025;28(1):254.
18. Martinez B, Pechlivanoglou P, Meng D, Traubici B, Mahood Q, Korczak D, et al. Clinical health outcomes of siblings of children with chronic conditions: a systematic review and meta-analysis. J Pediatr. 2022;250:83-92.e8.
19. Pinquart M. Behavior problems, self-esteem, and prosocial behavior in siblings of children with chronic physical health conditions: an updated meta-analysis. J Pediatr Psychol. 2023;48(1):77-90.

20. Wolff B, Magiati I, Roberts R, Pellicano E, Glasson EJ. Risk and resilience factors impacting the mental health and wellbeing of siblings of individuals with neurodevelopmental conditions: a mixed methods systematic review. Clin Psychol Rev. 2022;98:102217.

21. Powell TH, Ogle PA. Brothers and sisters: a special part of exceptional families. 3rd ed., Baltimore: Paul H. Brooks; 1985.

22. Valdivieso C, Ribley S, Ambler L. NICHCY New Digest; Children With Disabilities: Understanding sibling issues. Washington, Washington DS: Interstate Research Associates; 1988;11.

23. Bayhan P, Yükselen A. Sibling relationships of children who have disabled siblings. Journal of Selcuk University Faculty of Education. 2001;11:190-205.

24. Senel GH. Comparison of attitudes about insufficiency and their anxiety levels who have siblings with disability and those who do not. Journal of Special Education. 1995;2(1):33-39.

25. Ertürk B. (Advisor: Ahmetoğlu E.) Evaluation of the sibling relations of children with disabilities who participate in integrated education and those who do not, according to their mother and sibling perceptions Master's Thesis, Institute of Social Sciences, Trakya University. Department of Primary Education; Classroom Teaching. Edirne:2009.

26. Sloper P. Experiences and support needs of siblings of children with cancer. Health Soc Care Community. 2000;8(5):298-306.

27. Hanvey I, Malovic A, Ntontis E. Glass children: the lived experiences of siblings of people with a disability or chronic illness. J Community Appl Soc Psychol. 2022;32(5):936-948.

28. Bhattashali A, Ostrosky MM, Monda-Amaya L. Perceptions of typically developing children in India about their siblings with disabilities. International Journal of Inclusive Education. 2018;22(12):1257-1271.

29. Wawrzynski SE, Alderfer MA, Kvistad W, Linder L, Reblin M, Guo JW, et al. The social networks and social support of siblings of children with cancer. Children (Basel). 2022;9(1):113.

30. Le Boudec A, Akre C, Barrense-Dias Y, Berchtold A, Suris JC. Siblings of youths with chronic conditions: a school-based survey. Int J Adolesc Med Health. 2020;33(3):173-180.

31. Orm S, Haukeland Y, Vatne T, Silverman WK, Fjermestad K. Prosocial behavior is a relative strength in siblings of children with physical disabilities or autism spectrum disorder. Journal of Developmental and Physical Disabilities. 2022;34(4):591-608.

32. Shivers CM, Kozimor LM. Adolescent siblings of individuals with intellectual disabilities with and without comorbid mental health problems: a preliminary comparison of sibling perceptions. Journal of Mental Health Research in Intellectual Disabilities. 2017;10(1):1-16.

33. Shojaee S, Hemati Alamdarloo G, Nikoobin Borujeni F. Adjustment difficulties of siblings of children with disabilities and typically developing children. International Journal of Inclusive Education. 2020;24(4):414-426.

34. Tomeny TS, Baker LK, Barry TD, Eldred SW, Rankin JA. Emotional and behavioral functioning of typically-developing sisters of children with autism spectrum disorder: the roles of ASD severity, parental stress, and marital status. Research in Autism Spectrum Disorders. 2016;32:130-142.

35. Tomeny TS, Barry TD, Fair EC. Parentification of adult siblings of individuals with autism spectrum disorder: distress, sibling relationship attitudes, and the role of social support. Journal of Intellectual & Developmental Disability. 2016;42(4):320-331.

36. Tsai HJ, Cebula K, Liang SH, Fletcher-Watson S. Siblings' experiences of growing up with children with autism in Taiwan and the United Kingdom. Res Dev Disabil. 2018;83:206-216.

37. Shivers CM. Self-reported guilt among adult siblings of people with intellectual and developmental disabilities. Am J Intellect Dev Disabil. 2019;124(5):470-477.

38. Sommantico M, Parrello S, De Rosa B. Adult siblings of people with and without intellectual and developmental disabilities: sibling relationship attitudes and psychosocial outcomes. Res Dev Disabil. 2020;99:103594.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):185-190
DOI: 10.4274/BMB.galenos.2026.68077

The Relationship Between Plantar Fasciitis and Knee Osteoarthritis

Plantar Fasiit ve Diz Osteoartriti Arasındaki İlişki

Emin Can Balcı, Sever Çağlar, Abdurrahman Acar, Anıl Şüküroğlu, Serdar Biber

University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Orthopedics and Traumatology, İstanbul, Turkey

Abstract

Objective: In this study, we aimed to determine the prevalence and staging of gonarthrosis in patients diagnosed with plantar fasciitis. We also evaluated the coexistence of these two conditions, which may share common etiopathogenetic factors.

Method: A total of 171 participants, including 57 patients diagnosed with plantar fasciitis and 114 healthy controls, were included in the study. The diagnosis of plantar fasciitis in patients aged 30-70 was based on symptoms such as prominent heel pain during the first steps in the morning, tenderness upon palpation of the medial calcaneal tubercle, and a positive Windlass test, with radiological evaluation for the presence of calcaneal spurs. Individuals with lower extremity pathologies, previous surgical history, rheumatologic diseases, neurological or vascular causes of heel pain, and a body mass index ≥30 were excluded. The control group consisted of individuals without plantar fasciitis, with similar age and gender distribution. All participants underwent standing anteroposterior and lateral knee radiographs, and gonarthrosis staging was performed using the Kellgren-Lawrence classification.

Results: A total of 57 patients diagnosed with plantar fasciitis and 114 control individuals were included in the study. Upon examining the demographic data of the patients, no statistically significant differences were found between the groups. The frequency of radiological knee osteoarthritis was 77.2% in the patient group and 58.8% in the control group, with the patient group showing a significantly higher prevalence (p=0.017). Furthermore, when evaluating the stages of osteoarthritis, the proportion of advanced-stage osteoarthritis was significantly higher in the patient group compared to the control group (p=0.020).

Öz

Amaç: Bu çalışmada, plantar fasiit tanılı hastalarda gonartroz görülme oranını ve evrelemesini belirlemeyi amaçladık. Etiyopatogenezinde ortak faktörler bulunabilecek bu iki hastalığın birliktelik gösterme durumunu değerlendirdik.

Yöntem: Plantar fasiit tanısı olan 57 hasta ve 114 sağlıklı kontrol bireyi olmak üzere toplam 171 kişi çalışmaya dahil edildi. Otuz-70 yaş aralığındaki hastaların plantar fasiit tanısı, sabah ilk adımlarda belirginleşen topuk altı ağrısı, medial kalkaneal tübere palpasyonla hassasiyet ve Windlass testi pozitifliği gibi bulgulara dayanarak konulmuş, kalkaneal spur varlığı radyolojik olarak değerlendirilmiştir. Plantar fasiit dışında alt ekstremite patolojisi, geçirilmiş cerrahi öyküsü, romatolojik hastalık, nörolojik veya vasküler nedenli topuk ağrısı olanlar ve vücut kitle indeksi ≥30 bireyler dışlanmıştır. Kontrol grubunda, plantar fasiit tanısı olmayan, yaş ve cinsiyet dağılımı benzer bireyler yer almıştır. Tüm hastalara ayakta yük vererek diz ön-yan grafileri çekilmiş ve Kellgren-Lawrence sınıflaması ile gonartroz evrelemesi yapılarak gruplar arasında karşılaştırılmıştır.

Bulgular: Plantar fasiit tanısı olan 57 hasta ve 114 kontrol bireyi çalışmaya dahil edildi. Hasta demografik verileri incelendiğinde, yaş, cinsiyet dağılımı ve vücut kitle indeksi açısından gruplar arasında istatistiksel olarak anlamlı bir fark bulunmadı. Radyolojik diz osteoartrit sıklığı olgu grubunda %77,2, kontrol grubunda ise %58,8 olup, olgu grubunda anlamlı olarak daha yüksek bulundu (p=0,017). Ayrıca, artroz evreleri değerlendirildiğinde, olgu grubunda ileri evre artroz oranı kontrol grubuna göre anlamlı olarak daha yüksekti (p=0,020).

Address for Correspondence: Emin Can Balcı, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Orthopedics and Traumatology, İstanbul, Turkey
E-mail: emincanbalci@gmail.com **ORCID:** orcid.org/0000-0002-1565-9620
Received: 08.02.2025 **Accepted:** 17.06.2026 **Publication Date:** 24.06.2026

Cite this article as: Balcı EC, Çağlar S, Acar A, Şüküroğlu A, Biber S. The relationship between plantar fasciitis and knee osteoarthritis. Bagcilar Med Bull. 2026;11(2):185-190

Abstract

Conclusion: This finding suggests that knee osteoarthritis is more frequently observed in individuals diagnosed with plantar fasciitis. This result carries an important message for clinical practice, indicating that the assessment of knee joint symptoms in patients with plantar fasciitis should be considered, and, when necessary, early intervention may help improve quality of life and prevent progressive joint damage.

Keywords: Heel pain, knee osteoarthritis, plantar fasciitis

Öz

Sonuç: Plantar fasiit tanısı alan bireylerde diz osteoartritinin daha sık görüldüğünü göstermektedir. Bu bulgu, klinik pratiğe önemli bir mesaj taşımakta olup, plantar fasiitli hastaların değerlendirilmesinde diz eklemi semptomlarının dikkate alınmasının ve gerektiğinde erken müdahalenin yaşam kalitesini artırabileceği ve ilerleyici eklem hasarını önlemeye katkı sağlayabileceği düşünülmektedir.

Anahtar kelimeler: Diz osteoartriti, plantar fasiit, topuk ağrısı

Introduction

Plantar fasciitis is one of the most common causes of heel pain. It is a condition that becomes more pronounced with the first steps in the morning and negatively impacts daily living activities (1). The etiology of the disease involves excessive loading of the plantar fascia, microtrauma, and biomechanical disorders (2,3). Similarly, gonarthrosis (knee osteoarthritis) is a frequently encountered degenerative joint disease in the weight-bearing joints of the lower extremities. Gonarthrosis leads to pain, restricted mobility, and loss of function, significantly affecting the quality of life of patients (4,5).

Lower extremity alignment disorders may play a common role in the pathogenesis of both plantar fasciitis and gonarthrosis. Specifically, changes in mechanical load distribution, such as varus or valgus alignment, may increase the stress on the plantar fascia and knee joint. This may increase the likelihood of co-occurrence of both diseases. However, studies on the coexistence of plantar fasciitis and gonarthrosis in the literature are limited, and there is insufficient information regarding the common mechanisms of these two conditions.

In this study, we aimed to determine the prevalence of gonarthrosis in patients diagnosed with plantar fasciitis, evaluate the potential effects of lower extremity alignment disorders on the coexistence of these two diseases, and investigate the impact of the presence of gonarthrosis on the success of plantar fasciitis treatment.

Materials and Methods

Study Design

This study is a single-center retrospective comparative observational study comparing the prevalence of gonarthrosis in patients diagnosed with plantar fasciitis and a healthy control group. The study was conducted in accordance with the World Medical Association's Declaration of Helsinki. Patients were included from our medical practice between July 2024 and December 2024. After obtaining informed consent, patient data were extracted from our medical records and included in the study. The study approved by University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital Clinical Research Ethics Committee (approval number: 2024/03/11/035, date: 22.03.2024).

Patients who were definitively diagnosed with plantar fasciitis through clinical evaluation were included in the study. Individuals aged between 30 and 70 years were enrolled. The following typical examination findings were used to diagnose plantar fasciitis: Heel pain, particularly noticeable with the first steps in the morning or after standing up following prolonged sitting, tenderness and pain on palpation of the medial calcaneal tuberosity along the plantar fascia, and increased pain during dorsiflexion of the toes (positive Windlass test) (2,3). When necessary, radiological evaluation was used to support the diagnosis of plantar fasciitis, considering the presence of a calcaneal spur (heel spur) (Figure 1). Patients with additional foot or lower extremity pathologies other than plantar fasciitis were excluded from this study. Furthermore, individuals with congenital or acquired foot deformities, those with a history of lower extremity surgery, and those with a rheumatological disease were not included in the study. Additionally, patients with heel pain of neurological or vascular origin and individuals without typical examination findings supporting the diagnosis of plantar fasciitis were also excluded. Lastly, individuals with a body mass index above 30 were excluded from the study. These criteria were established to ensure the homogeneity of the study population, minimize the effects of confounding factors, and enhance the reliability of the results.

The control group consisted of individuals without a diagnosis of plantar fasciitis who were matched to the plantar fasciitis group in terms of age and sex distribution.

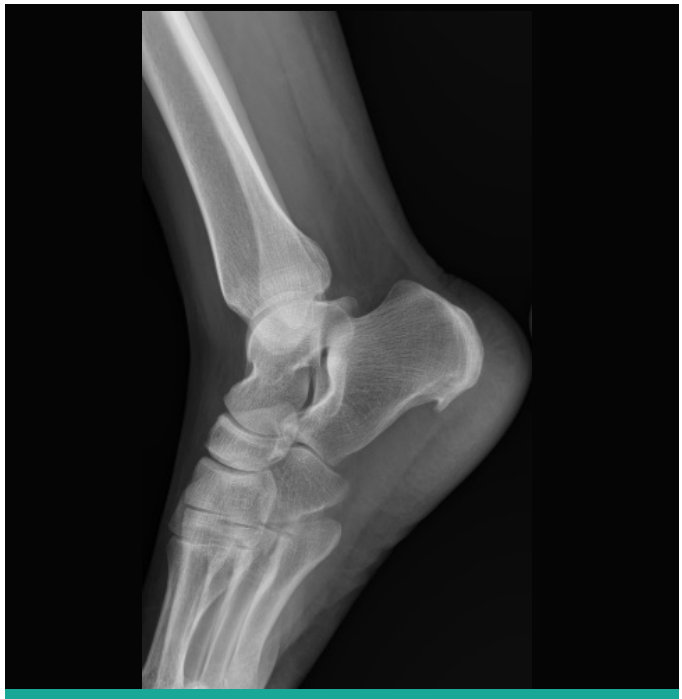


Figure 1. Calcaneal spur

In this study, patients diagnosed with plantar fasciitis in the outpatient clinic were evaluated for the presence of knee osteoarthritis using knee radiographs that were already available in their medical records. These knee radiographs were obtained concurrently during routine foot and ankle radiographic examinations performed to evaluate calcaneal spur, with the knee included within the cassette; therefore, no additional radiation exposure was required for the purposes of this study.

The control group was selected from patients who presented to the emergency department with minor knee trauma and underwent knee radiography for clinical indications. Individuals with acute fractures, dislocations, inflammatory joint disease, previous knee surgery, or chronic knee pain were excluded. The control group was not selected according to the presence or absence of radiographic knee osteoarthritis. All patients had provided informed consent for the use of their medical records for research purposes.

Radiographic Classification of Knee Osteoarthritis

Patients were evaluated through weight-bearing anteroposterior and lateral knee radiographs, and the presence of osteoarthritis was determined using the Kellgren-Lawrence classification. The Kellgren-Lawrence classification is a widely used method for grading knee osteoarthritis and consists of four stages. Stage 0 indicates the absence of osteoarthritis findings, with no signs of joint disease observed. Stage 1 is characterized by minimal osteoarthritic changes, typically showing small osteophytes or mild cartilage thinning. Stage 2 is associated with moderate osteoarthritis, with prominent findings such as osteophytes, cartilage thinning, joint space narrowing, and subchondral sclerosis. Stage 3 indicates moderate osteoarthritis, characterized by multiple osteophytes, definite joint space narrowing, and subchondral sclerosis. Stage 4 represents severe osteoarthritis, with marked joint space narrowing, large osteophytes, severe subchondral sclerosis, and bony deformity. Due to the limited number of patients with Kellgren-Lawrence grade 4 osteoarthritis, stages 3 and 4 were combined for statistical analysis (Figure 2) (6). Based on the X-ray findings, the osteoarthritis stage for each patient's knee was determined, and these findings were compared between the plantar fasciitis group and the control group.

Sample Size and Matching

Power analysis was conducted based on an average effect size of 0.5 calculated across all measures. As a result of this analysis, a sample size of 80 participants was determined, consisting of 40 patients diagnosed with plantar fasciitis and 40 individuals without a diagnosis of plantar fasciitis, forming a matched control group. This sample size was chosen to provide sufficient statistical power to detect meaningful differences between the groups. Matching the patient and control groups based on fundamental demographic factors such as age and gender enhances the robustness of the study. By using this sample size, the goal is to obtain reliable and generalizable findings regarding the relationship between plantar fasciitis and knee osteoarthritis.

Statistical Analysis

Descriptive statistics for the data included mean, standard deviation, median, minimum, maximum, frequency, and percentage values. The distribution of variables was assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests. The Mann-Whitney U test was used for the analysis of independent quantitative data. The Wilcoxon test was used for the analysis of dependent quantitative data. The chi-square test was employed for the analysis of independent categorical data. All analyses were performed using SPSS 27.0 software.

Results

A total of 57 patients diagnosed with plantar fasciitis and 114 control patients were included in the study. The

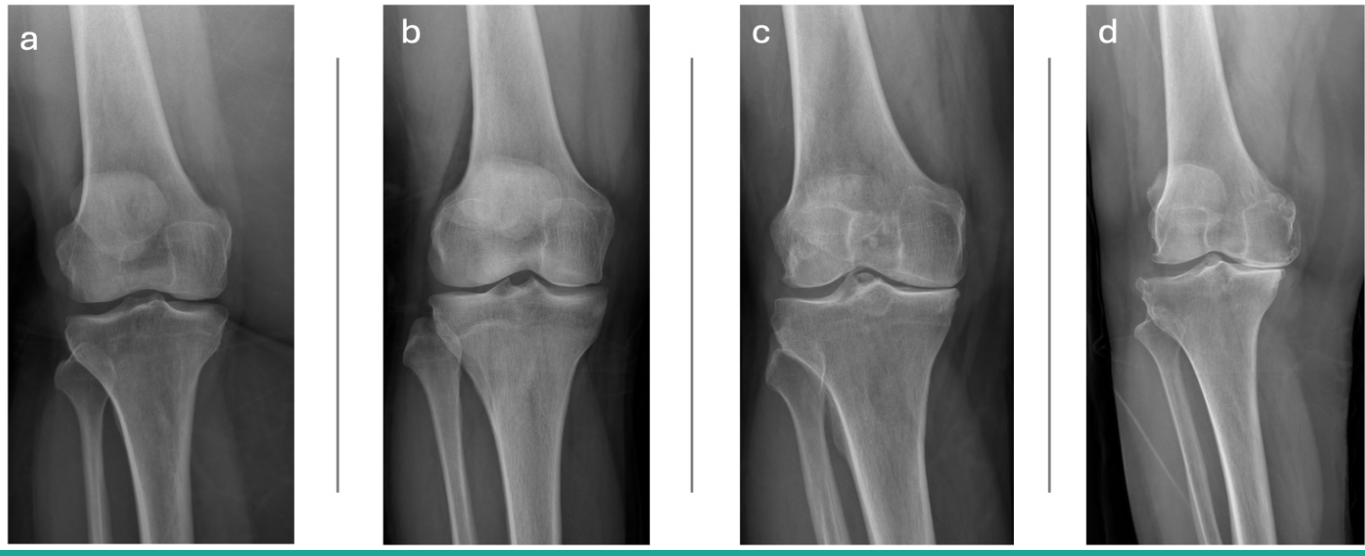


Figure 2. Kellgren-Lawrence classification (a) Stage 0 (b) Stage 1 (c) Stage 2 (d) Stage 3

demographic information of the patients is reported in Table 1. The demographic data did not show any statistically significant differences and exhibited a normal distribution.

The presence and staging of osteoarthritis in the patient and control groups are presented in Table 2.

The osteoarthritis prevalence was significantly higher in the case group compared to the control group (p=0.017) (Table 2, Figure 3). Additionally, the osteoarthritis stage was significantly higher in the case group than in the control group (p=0.020) (Table 2, Figure 4).

Discussion

In this study, we investigated the prevalence of gonarthrosis in patients diagnosed with plantar fasciitis, and the results revealed that radiographic knee osteoarthritis was more frequently observed among individuals with plantar fasciitis compared with the control group. This finding suggests that there may be an association between plantar fasciitis and gonarthrosis.

Table 1. Characteristics of the study group (mean ± standard deviation)

	Patient group (n=57)	Control group (n=114)
Mean age, y	46.7±10.6	43.2±12.3
Sex Female Male	 46 11	 81 33
Mean BMI (kg)	24.1±3.3	24.9±2.4

BMI: Body mass index

Table 2. Presence and staging of osteoarthritis (mean, percentage)

	Patient group (n=57)	Control group (n=114)
Osteoarthritis		
(-)	13 22.8%	47 41.2%
(+)	44 77.2%	67 58.8%
Stage		
0	13 22.8%	47 41.2%
1	6 10.5%	20 17.5%
2	22 38.6%	32 28.1%
3 & 4	16 28.1%	15 13.2%

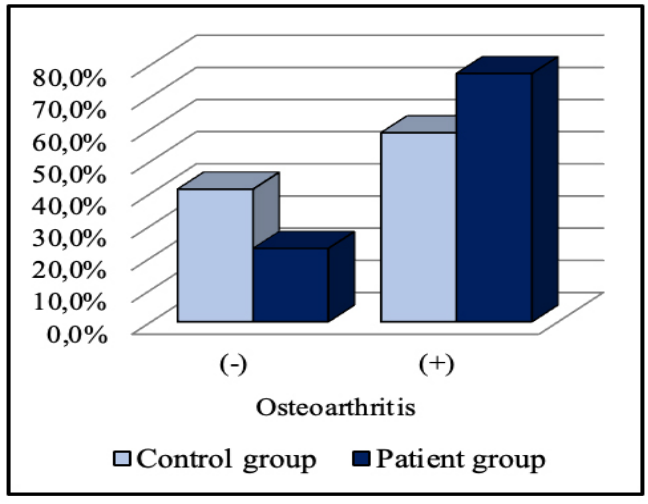


Figure 3. Osteoarthritis prevalence in the control and case groups

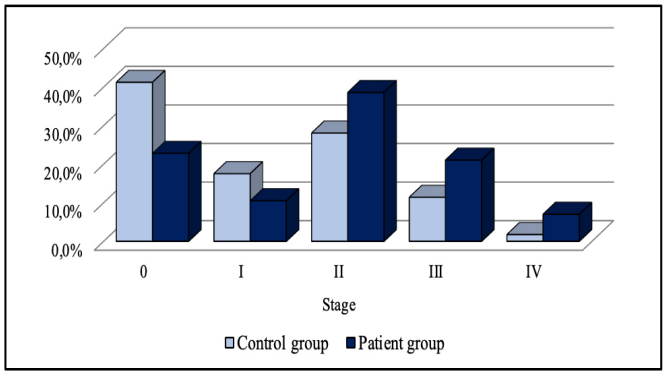


Figure 4. Osteoarthritis Stage distribution in the control and case groups

Similarly, the literature indicates that plantar fasciitis is frequently observed in patients with knee osteoarthritis, and that restricted ankle dorsiflexion is a significant risk factor for the development of plantar fasciitis (7). Our findings support this study. It is believed that restricted ankle dorsiflexion may increase biomechanical loading, creating excessive tension on the plantar fascia, which could play a role in the development of plantar fasciitis (7). In this context, rehabilitation programs aimed at increasing ankle range of motion may be important for preventing the development of plantar fasciitis or supporting the treatment process in patients with knee osteoarthritis. In a study where the plantar fascia was evaluated using ultrasonography in patients diagnosed with gonarthrosis, findings consistent with plantar fasciitis, such as thickening of the plantar fascia and increased echogenicity, were observed (8). Additionally, obesity is known to be a significant risk factor for both the development of plantar fasciitis and gonarthrosis (9). In our study, we excluded obese patients from the study groups, thereby preventing this risk factor from influencing the results of the study.

It has been shown that foot pathologies lead to symptoms that alter individuals' walking patterns, and these biomechanical changes increase the risk of developing knee osteoarthritis (10). Pronation or abnormal flatfoot gait is associated with many painful foot conditions, including plantar fasciitis (11-13). This condition has been suggested to increase rotational stress in the knee joint, according to several studies (14). The reason for this is the tight biomechanical connection between the movement of the rear foot and the tibia (15). As a result, abnormal movements in the foot can affect the knee joint. Over time, this abnormal stress may lead to the development of osteoarthritis in the knee joint. Specifically, a foot that is pronated during walking has been shown to be associated with knee pain and medial gonarthrosis of the knee (14). In patients diagnosed with plantar fasciitis, a shortened stance phase has been shown to affect knee biomechanics (16). In individuals with gonarthrosis, foot symptoms have been noted to increase the risk of worsening knee pain over the following 4 years, although they were not associated with the progression of osteoarthritis. These results are thought to be due to the short follow-up period in the related study for osteoarthritis progression (17).

The possible mechanisms underlying the high prevalence observed in our study need to be addressed. In individuals with plantar fasciitis, pain and functional loss in the sole of the foot can lead to changes in walking patterns. Such compensatory mechanisms may increase the load on the knee joint and contribute to the development of gonarthrosis in the long term. Additionally, the sedentary lifestyle commonly observed in individuals with plantar fasciitis may also increase the risk of gonarthrosis. There are a limited number of studies in the literature regarding the relationship between plantar fasciitis and gonarthrosis. However, our findings contribute to the literature by suggesting a potential link between these two conditions. Our study emphasizes the need for further research to explore this relationship in greater detail. In particular, prospective cohort studies and biomechanical analyses are crucial for better understanding the causality of this relationship.

Study Limitations

This study has several limitations. Gonarthrosis was assessed radiologically, but symptomatic scoring of the patients was not performed. Plantar fasciitis can cause symptoms of varying severity in patients, and the intensity of these symptoms may affect their relationship with gonarthrosis. No classification of symptom severity was made in this study. Future prospective cohort studies with larger groups will be valuable in guiding further research in this regard.

Conclusion

In conclusion, this study demonstrated an association between plantar fasciitis and a higher prevalence of radiographic knee osteoarthritis. Because of the retrospective observational design, these findings should be interpreted as an association rather than a causal relationship. This finding carries an important message for clinical practice. Attention to knee joint symptoms during

the evaluation of patients with plantar fasciitis and, if necessary, taking appropriate early measures, can improve quality of life and prevent progressive joint damage.

Ethics

Ethics Committee Approval: The study approved by University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital Clinical Research Ethics Committee (approval number: 2024/03/11/035, date: 22.03.2024).

Informed Consent: Written informed consent was obtained from all participants who participated in this study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: E.C.B., A.Ş., S.B., Concept: E.C.B., S.Ç., A.A., Design: E.C.B., S.Ç., A.A., Data Collection or Processing: E.C.B., A.Ş., S.B., Analysis or Interpretation: E.C.B., A.A., A.Ş., S.B., Literature Search: E.C.B., S.Ç., A.A., Writing: E.C.B.

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

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

References

1. Trojian T, Tucker AK. Plantar fasciitis. Am Fam Physician. 2019;99(12):744-750.
2. Rhim HC, Kwon J, Park J, Borg-Stein J, Tenforde AS. A systematic review of systematic reviews on the epidemiology, evaluation, and treatment of plantar fasciitis. Life. 2021;11(12):1287.
3. Thompson JV, Saini SS, Reb CW, Daniel JN. Diagnosis and management of plantar fasciitis. J Am Osteopath Assoc. 2014;114(12):900-906.
4. Lespasio MJ, Piuzzi NS, Husni ME, Muschler GF, Guarino A, Mont MA. Knee osteoarthritis: a primer. TPJ. 2017;21(4):16-183.
5. Zhang Y, Jordan JM. Epidemiology of osteoarthritis. Clin Geriatr Med. 2010;26(3):355-369.
6. Kohn MD, Sassoon AA, Fernando ND. Classifications in brief: Kellgren-Lawrence classification of osteoarthritis. Clin Orthop Relat Res. 2016;474(8):1886-1893.
7. Miladi S, Bouzid S, Fazaa A, Boussaa H, Makhlouf Y, Souabni L, et al. Is there an association between plantar fasciitis and knee osteoarthritis? Musculoskeletal Care. 2023;21(4):1045-1052.
8. Liu Z, Sui X, Hui R, Zhao Y, Li H, Huang X. Ultrasonic evaluation of plantar fascia in patients with osteoarthritis of the knee. Medicine (Baltimore). 2022;101(38):e30613.
9. Frey C, Zamora J. The effects of obesity on orthopaedic foot and ankle pathology. Foot Ankle Int. 2007;28(9):996-999.
10. Mickle KJ, Munro BJ, Lord SR, Menz HB, Steele JR. Cross-sectional analysis of foot function, functional ability, and health-related quality of life in older people with disabling foot pain. Arthritis Care Res (Hoboken). 2011;63(11):1592-1598.
11. Golightly YM, Hannan MT, Dufour AB, Hillstrom HJ, Jordan JM. Foot disorders associated with overpronated and oversupinated foot function: the Johnston County osteoarthritis project. Foot Ankle Int. 2014;35(11):1159-1165.
12. Irving DB, Cook JL, Young MA, Menz HB. Obesity and pronated foot type may increase the risk of chronic plantar heel pain: a matched case-control study. BMC Musculoskelet Disord. 2007;8:41.
13. Menz HB, Dufour AB, Riskowski JL, Hillstrom HJ, Hannan MT. Association of planus foot posture and pronated foot function with foot pain: the Framingham foot study. Arthritis Care Res (Hoboken). 2013;65(12):1991-1999.
14. Gross KD, Felson DT, Niu J, Hunter DJ, Guermazi A, Roemer FW, et al. Association of flat feet with knee pain and cartilage damage in older adults. Arthritis Care Res (Hoboken). 2011;63(7):937-944.
15. Souza TR, Pinto RZ, Trede RG, Kirkwood RN, Fonseca ST. Temporal couplings between rearfoot-shank complex and hip joint during walking. Clin Biomech (Bristol). 2010;25(7):745-758.
16. Riskowski JL, Hagedorn TJ, Dufour AB, Hannan MT. Associations of region-specific foot pain and foot biomechanics: the Framingham foot study. J Gerontol A Biol Sci Med Sci. 2015;70(10):1281-1288.
17. Paterson KL, Kasza J, Hunter DJ, Hinman RS, Menz HB, Peat G, et al. Longitudinal association between foot and ankle symptoms and worsening of symptomatic radiographic knee osteoarthritis: data from the osteoarthritis initiative. Osteoarthritis Cartilage. 2017;25(9):1407-1413.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):191-197
DOI: 10.4274/BMB.galenos.2026.94695

Examination of Hematological Parameters after Carbon Filtered Hemoperfusion in Patients with Mushroom Poisoning

Mantar Zehirlenmesi olan Hastalarda Karbon Filtreli Hemoperfüzyon Sonrası Hematolojik Parametrelerin İncelenmesi

Mehmet Burak Durlanık[1], Serhat Mert Tiril[2], Rıdvan Sivritepe[3], Sema Basat[2]

[1]Akyazı State Hospital, Clinic of Internal Medicine, Sakarya, Turkey
[2]University of Health Sciences Turkey, Ümraniye Training and Research Hospital, Department of Internal Medicine, İstanbul, Turkey
[3]İstanbul Medipol University Faculty of Medicine, Department of Internal Medicine, İstanbul, Turkey

Abstract

Objective: It is known that hemoperfusion with carbon filter has important benefits in the treatment of mushroom poisoning. In this study, the changes in hematological values after hemoperfusion with carbon filter in patients with mushroom poisoning were investigated.

Method: This study was planned as a retrospective study. Fifty-one patients who were hospitalized in the internal medicine service due to mushroom poisoning and underwent hemoperfusion with carbon filter between 01.01.2015 and 31.12.2021 in our hospital were included in the study. The patients were evaluated in terms of age, gender, time to onset of complaints after mushroom ingestion, time to start of hemoperfusion after mushroom ingestion, comorbid diseases, type of mushroom eaten, complaints at the time of admission, duration of hospitalization, daily hematological and biochemical parameters before and after carbon filter hemoperfusion, the lowest platelet value of the patients after carbon filter hemoperfusion and the day on which this value was observed.

Results: The mean age of the patients included in the study was 48.29 years. It was observed that 51.0% of the patients were male and 49.0% were female. Among the comorbidity distributions of the patients, hypertension (25.5%) was found the most and chronic renal failure (2.0%) the least. Among the complaints, the most nausea (76.5%) and the least abdominal pain (25.5%) were detected. When the mushroom type was examined, it was determined that

Öz

Amaç: Mantar zehirlenmesi tedavisinde karbon filtreli hemoperfüzyon yapılmasının önemli yararları olduğu bilinmektedir. Bu çalışmada mantar zehirlenmesi olan hastalarda karbon filtreli hemoperfüzyon sonrası hematolojik değerlerinin değişimi incelenmiştir.

Yöntem: Bu çalışma retrospektif bir çalışma olarak planlanmıştır. Hastanemizde 01.01.2015-31.12.2021 tarihleri arasında mantar zehirlenmesi nedeniyle iç hastalıkları servisine yatırılan ve karbon filtreli hemoperfüzyon yapılan 51 hasta çalışmaya alınmıştır. Hastalar; yaş, cinsiyet, mantar yenmesi sonrasında şikayetlerin başlama süresi, mantar yenmesi sonrasında hemoperfüzyona başlanma süresi, komorbid hastalıklar, yenilen mantarın cinsi, başvuru sırasındaki yakınmaları, hastaların hastanede yatış süresi, karbon filtreli hemoperfüzyon öncesi ve sonrasında günlük hematolojik ve biyokimyasal parametreleri, hastaların karbon filtreli hemoperfüzyon sonrası en düşük trombosit değeri ve bu değerin görüldüğü gün açısından değerlendirildi.

Bulgular: Çalışmaya dahil edilen hastaların yaş ortalaması 48,29 yıl idi. Hastaların %51,0'ı erkek, %49,0'ı kadın olduğu görülmüştür. Hastaların komorbidite dağılımları arasında en fazla hipertansiyon (%25,5), en az kronik böbrek yetmezliği (%2,0) olarak tespit edilmiştir. Şikayetler arasında en fazla bulantı (%76,5), en az karın ağrısı (%25,5) tespit edilmiştir. Mantar türü incelendiğinde ise toplama mantar %88,2, kültür mantarı %11,8 olarak tespit edilmiştir. Karbon filtreli

Address for Correspondence: Assoc. Prof., Rıdvan Sivritepe, İstanbul Medipol University Faculty of Medicine, Department of Internal Medicine, İstanbul, Turkey
E-mail: ridvan.sivritepe@medipol.edu.tr **ORCID:** orcid.org/0000-0003-0547-1883
Received: 02.03.2025 **Accepted:** 25.02.2026 **Epub:** 04.03.2026 **Publication Date:** 24.06.2026

Cite this article as: Durlanık MB, Tiril SM, Sivritepe R, Basat S. Examination of hematological parameters after carbon filtered hemoperfusion in patients with mushroom poisoning. Bagcilar Med Bull. 2026;11(2):191-197

Abstract

the foraged mushrooms was 88.2% and the cultivated mushrooms was 11.8%. When the carbon filtered hemoperfusion onset time was examined, it was found that 45.1% of the patients started early and 54.9% of the patients started later. The average hospitalization day was determined as 7.18±4.16 days. It is seen that there is a statistically significant difference between the values of 0-1 day, 0-2 days, 0-3 days, 1-2 days and 1-3 days for the thrombocyte value. The highest thrombocyte value (236 10^3/μL) was seen on day 0 ($p<0.001$). After the first carbon filtered hemoperfusion, the mean values of the patients with the lowest thrombocyte value were 58.51 10^3/μL, and the average of the days when this value was seen was 2.5 days.

Conclusion: When hematological parameters were examined after hemoperfusion with carbon filter in patients with mushroom poisoning, thrombocyte value decreased in all patients and increased to normal values within days without causing complications.

Keywords: Carbon filtered hemoperfusion, mushroom poisoning, thrombocytopenia

Öz

hemoperfüzyon başlama zamanı incelendiğinde %45,1 hastada erken, %54,9 hastada daha geç başlandığı tespit edilmiştir. Yatış gün ortalaması 7,18±4 gün olarak tespit edilmiştir. Trombosit değeri için 0-1 gün, 0-2 gün, 0-3 gün, 1-2 gün, 1-3 gün değerleri arasında istatistiksel olarak anlamlı fark olduğu görülmektedir. En yüksek trombosit değeri (236 10^3/μL) 0. gün görülmüştür ($p<0,001$). Hastaların ilk karbon filtreli hemoperfüzyon sonrası en düşük trombosit değerinin görüldüğü değerlerin ortalaması 58,51 10^3/μL ve bu değerin görüldüğü günlerin ortalaması 2,5 gün olarak tespit edilmiştir.

Sonuç: Mantar zehirlenmesi olan hastalarda karbon filtreli hemoperfüzyon sonrası hematolojik parametreler incelendiğinde hastaların tümünde trombosit değerinde düşüş gözlenmiştir ve komplikasyona neden olmadan günler içerisinde normal değerlere yükselmiştir.

Anahtar kelimeler: Karbon filtreli hemoperfüzyon, mantar zehirlenmesi, trombositopeni

Introduction

The disease caused by some poisonous compounds contained in capped mushrooms is called mushroom poisoning. There are approximately 5.000 species of mushrooms in nature, and approximately 100 of them contain toxins (1). As a result of ingestion of wild mushroom species, the clinical picture of the patient varies from gastrointestinal system complaints to severe cytotoxic conditions that can result in multiorgan failure and death (2).

The most lethal mushroom poisoning occurs with Amanita (Amanita phalloides, Amanita verna, Amanita ocreata) and Galerina species, which contain cyclopeptides. Amanita phalloides is responsible for 50% of all mushroom poisonings, and all Amanita species are responsible for 95% of total mortality (3).

According to recent epidemiological studies, mushroom poisoning accounts for approximately 3-5% of all acute poisoning cases in Turkey, with the majority caused by foraged wild mushrooms. The incidence shows seasonal variability, with peaks in spring and autumn, especially in rural and forested regions where uncontrolled mushroom consumption is prevalent (4,5).

Hemoperfusion is the direct passage of blood through various adsorbents such as charcoal or resin and is used to remove toxic compounds from the circulatory system. Hemoperfusion has been used experimentally or therapeutically over the past several years to treat uremia, poisoning and drug intoxication, liver failure, and to increase the clearance of a variety of other noxious substances (6,7).

Carbon-filtered hemoperfusion is increasingly considered as an effective extracorporeal detoxification method due to its ability to adsorb protein-bound and lipid-soluble toxins such as amatoxins. Its rapid toxin clearance profile offers a clinical advantage over conventional dialysis methods in cases of mushroom poisoning with delayed symptom onset (8).

This study aimed to retrospectively investigate the changes in hematological parameters of patients who were hospitalized in the Department of Internal Medicine, University of Health Sciences Turkey, Ümraniye Training and Research Hospital due to mushroom poisoning and underwent carbon filter hemoperfusion between 2015 and 2021.

Materials and Methods

In this study, 51 patients who were hospitalized in the Department of Internal Medicine, University of Health Sciences Turkey, Ümraniye Training and Research Hospital with the diagnosis of mushroom poisoning and underwent hemoperfusion with a carbon filter between 01.01.2015 and 31.12.2021 were evaluated retrospectively. The patients were evaluated in terms of age, gender, time to onset of complaints after mushroom ingestion, time to start of hemoperfusion after mushroom ingestion, comorbid diseases, type of mushroom eaten, complaints at the time of admission, duration of hospitalization, daily hematological and biochemical parameters [leukocyte,

hemoglobin, platelet, mean platelet volume (MPV), alanine aminotransferase, aspartate aminotransferase, alkaline phosphatase, Gamma glutamyl transferase (GGT), lactate dehydrogenase, total bilirubin, international normalized ratio, sodium, potassium, blood urea nitrogen, creatinine] for three days before and after carbon filter hemoperfusion, the lowest platelet value of the patients after carbon filter hemoperfusion and the day on which this value was observed.

The patients were first divided into three groups according to their ages: 15-25, 26-65 and over 65. The time of onset of symptoms after mushroom ingestion was then divided into two groups: 0-6 hours and over 6 hours. The time of starting carbon filter hemoperfusion after mushroom ingestion was divided into two groups as 0-20 hours and over 21 hours. Comorbid diseases were evaluated as the presence or absence of hypertension, diabetes mellitus, congestive heart failure, chronic renal failure, and asthma in patients. The type of mushrooms was divided into two groups as foraged mushrooms and cultivated mushrooms. Complaints at the time of admission were evaluated as nausea, vomiting, diarrhea and abdominal pain. The length of stay of the patients was divided into three groups: 0-5, 6-10, and 10 days and above.

The patients included in the study did not have any known liver disease. Patients who were hospitalized due to mushroom poisoning and did not undergo hemoperfusion with a carbon filter were not included in the study. Exclusion criteria included known chronic liver disease, hematologic disorders, pre-existing thrombocytopenia, or any alternative diagnosis explaining the clinical presentation. All included patients had documented ingestion of wild mushrooms, symptom onset within 24 hours, and received hemoperfusion within the first 48 hours of admission. Patient information was obtained from outpatient clinic records and hospitalization files. The Declaration of Helsinki, which regulates biomedical research on humans, was followed. Additionally, Local Ethics Committee approval of University of Health Sciences Turkey, Ümraniye Training and Research Hospital (no: B.10.1.TKH.4.34.H.GP.0.01/363, date: 16.12.2021) was obtained. Informed consent was waived due to the retrospective nature of the study.

Statistical Analysis

Data were analyzed using the SPSS 25.0 package program. The Kolmogorov-Smirnov test was used to check whether the distribution of the data was normally distributed. While evaluating the study data, descriptive statistical methods such as median, minimum, maximum and ratio values were used. Wilcoxon test was used for comparison of two dependent groups with non-parametric distribution and Friedman test was used for comparison of more than two dependent groups. Significance was evaluated at $p<0.05$ levels for all values.

Results

The mean age of the patients included in the study was determined as 48.29±18.05 years. The distribution of age groups was found to be 19.6% (n=10) in the 15-25 age group, 54.9% (n=28) in the 25-65 age group, and 25.5% (n=13) in the 65 and above age group. Of the patients included in the study, 51.0% were male and 49.0% were female. Among the comorbidity distributions of the patients, the highest rate was hypertension 25.5%, and the lowest rate was chronic renal failure 2.0%. Among the complaints of the patients, the most common was nausea (76.5%) and the least common was abdominal pain (25.5%). When the mushroom species were examined, it was determined that the foraged mushrooms were 88.2% and the cultivated mushrooms were 11.8%. When symptom onset times were examined, symptoms began within the first 0-6 hours in 68.6% of patients and over 6 hours in 31.4%. When the starting times of carbon filter hemoperfusion were examined, It was observed that 45.1% of the patients were dialysed within the first 0-20 hours, and 54.9% after 21 hours. While the average length of stay was determined as 7.18±4.16 days, 39.2% of the patients were hospitalized for 0-5 days, 47.1% for 6-10 days, and 13.7% for more than 11 days (Table 1).

Laboratory values of the patients were divided into 4 groups: at the time of admission, day 1, day 2 and day 3. Wilcoxon test was used for pairwise comparison between days to detect differences. In the comparison between days, it is seen that there is a statistically significant difference in alkaline phosphatase, alanine aminotransferase, aspartate aminotransferase, blood urea nitrogen, creatinine, hemoglobin, potassium and lactate dehydrogenase values and the highest value is day 0 for all 8 parameters ($p<0.005$ for all). When the parameters of the patients were compared between days, a statistically significant difference was found between the international normalized ratio, GGT, MPV, and sodium values of the patients. It was observed that the highest value for GGT value was on the 3rd day, the highest value for MPV value was on the 2nd and 3rd days, and the highest value for Na value was on the 3rd day ($p<0.005$ for all).

When the platelet value is compared for days 0-1, days 0-2, days 0-3, days 1-2, and days 1-3, it is seen that there is a

statistically significant difference, and the highest value is day 0 (p<0.001). The mean of the lowest platelet value after the first carbon filter hemoperfusion was 58.51 10^3/µL and the mean of the days when this value was seen was 2.5 days. The patients' platelet values returned to normal after an average of 6.62 days, and the average platelet count on these days was found to be 198.7 10^3/µL (Table 2).

Discussion

In this study, we aimed to observe the changes in hematological and biochemical parameters after the application of carbon filter hemoperfusion in patients hospitalized due to mushroom poisoning. In the present study, we observed a decrease in the platelet count of patients after hemoperfusion with a carbon filter. It was

Table 1. Clinical and demographic data of patients

Age (years)	48.29±18.05		
Age group analysis	<25	25-65	>65
	10 (19.6%)	28 (54.9%)	13 (25.5%)
Sex	**Male**		**Female**
	26 (51%)		25 (49%)
Time to symptom onset	0-6 hours		6 hours <
	35 (68.6%)		16 (31.4%)
Type of mushroom	Foraged mushrooms		Cultivated mushroom
	45 (88.2%)		6 (11.8%)
Starting time for dialysis with carbon filter	0-20 hours		20 hours <
	23 (45.1%)		28 (54.9%)
Length of stay	0-5 days	6-10 days	10 days >
	20 (39.2%)	24 (47.2%)	7 (13.7%)
Comorbidity	Hypertension		Chronic renal failure
	13 (25.55)		1 (2%)
Complaints	Nausea		Abdominal pain
	39 (76.5%)		13 (25.5%)

Table 2. Comparison of laboratory values of patients by day

	Moment of admission	**The first day**	**The second day**	**The third day**	
	Median (min-max)	**Median (min-max)**	**Median (min-max)**	**Median (min-max)**	**p-value**
Alkaline phosphatase (38-155 U/L)	67 (30-371)	60 (28-379)	55 (29-326)	53 (24-273)	**p<0.001***
Alanine aminotransferase (10-40 U/L)	24 (2-990)	22 (6-4000)	20 (6-3242)	22 (5-2648)	**0.011****
Aspartate aminotransferase (15-50 IU/L)	28 (11-1247)	21 (11-5516)	20 (9-2525)	22 (9-710)	**p<0.001***
Blood urea nitrogen (10-20 mg/dL)	32 (14-111)	29 (12-111)	26 (8-107)	23 (7.6-104)	**p<0.001***
Creatinine (<1 mg/dL)	0.82 (0.57-2.9)	0.74 (0.47-4.02)	0.73 (0.48-3.5)	0.72 (0.44-4.59)	**0.001****
Gamma glutamyl transferase (7-49 U/L)	26 (8-700)	23 (7-574)	24 (7-450)	30 (7-397)	**p<0.001***
Hemoglobin (12.4-14.8 g/L)	14.4 (9.8-17.8)	12.5 (8.9-15.9)	11.6 (8.17-14.8)	11.1 (6.5-15.4)	**p<0.001***
International normalized ratio	1.11 (0.89-3.95)	1.25 (0.96-5.55)	1.2 (0.99-2.95)	1.17 (0.95-3.94)	**p<0.001***
Potassium (3.5-5.5 mmol/L)	4.4 (3-5.8)	4.1 (3-5.2)	3.9 (2.7-5.4)	4 (3.2-5.1)	**p<0.001***
Lactate dehydrogenase (90-250 U/L)	322 (161-2801)	241 (134-7317)	254 (129-2118)	240 (116-1254)	**p<0.001***
Mean platelet volume	9.5 (6.4-13.2)	9.9 (6.67-13.4)	10.2 (6.4-88.9)	10.2 (7.1-13.7)	**0.003****
Sodium (135-145 mEq/L)	139 (130-145)	139 (130-143)	138 (13-145)	139 (131-144)	**0.034****
Thrombocyte (150-450.000/mm^3)	236 (112-422)	106 (35-284)	73 (17-135)	71 (18-181)	**p<0.001***
Total bilirubin (0.3-1.9 mg/dL)	0.7 (0.17-3.45)	0.97 (0.18-4.5)	0.69 (0.09-3.95)	0.61 (0.09-5)	**0.001****
Leukocyte (4.1-8.9 10^3/uL)	10.11 (4.77-23.09)	10.2 (2.45-22.7)	8.45 (0.56-16.18)	7.2 (2.47-16.3)	**p<0.001***

Friedman test *: p<0.001, **: p<0.05, min-max: Minimum-maximum

observed that the thrombocytopenia returned to normal without any complications in the following days.

It is known that mushrooms are living organisms that do not carry chlorophyll, live as parasites or saprophytes, and reproduce by spores. Spores are dispersed into the environment by wind and germinate under suitable climatic conditions. That's why edible and poisonous mushrooms can grow side by side (9). Turkey is rich in mushroom flora due to its suitable ecological conditions (10). In Turkey, the habit of collecting and eating mushrooms from forests or meadows, especially in the spring months, is quite common among people with low socio-economic status. For this reason, the number of mushroom poisoning cases resulting in death is considerably high in our country. It is also known that the poisonous content of raw or cooked mushrooms does not change (10,11). In small mushrooms, symptoms appear approximately three hours after ingestion. In larger mushrooms, including Amanita phalloides, symptoms appear within 6-24 hours after ingestion. According to the onset time of symptoms, those starting up to six hours are defined as early syndromes, and those starting after six hours are defined as late syndromes. Those with late-onset symptoms are 90-95% fatal (1). In the study it was determined that approximately 48% of the mushroom species causing poisoning were Amanita phalloides and approximately 52% were non-amanita species (1). Since toxicological examination could not be performed in our hospital, in our study the diagnosis of poisoning was made based on anamnesis and clinical findings. When the mushroom species were examined in our study, we found that 88.2% of the mushrooms were foraged mushrooms and 11.8% were cultivated mushrooms. In our study, it was observed that mushrooms collected from nature and eaten caused more mushroom poisoning.

The most common symptoms in patients presenting to the emergency department due to mushroom poisoning are nausea, vomiting, abdominal pain, diarrhea, agitation, loss of consciousness and encephalopathy (12,13) In the studies 79.5% of the patients had vomiting and diarrhea, 10.3% had only vomiting, 5.1% had abdominal pain, and 5.1% had abdominal pain and vomiting at the time of admission (1,3,4). In the present study, complaints at the time of admission were nausea 76.5% (n=39), vomiting 72.5% (n=37), diarrhea 37.3% (n=19), and abdominal pain 25.5% (n=13). In our study, like other studies, we observed that patients with mushroom poisoning had more frequent complaints of nausea and vomiting.

In the study conducted by Colak et al. (14), patients were divided into two groups according to the time of onset of symptoms. In 62.1% of the patients, symptoms appeared within the first 0-5 hours, and in 37.9%, symptoms appeared within 6-24 hours. In our study, symptoms started in the first 6 hours in 68.6% of the patients and after 6 hours in 31.4%. In our study, similar to other studies, we observed that symptoms of mushroom poisoning occur more frequently within the first 6 hours.

In our study, the average hospitalization day was determined as 7.18±4.16. 39.2% of the patients were hospitalized for 0-5 days, 47.1% for 6-10 days, and 13.7% for more than 11 days. In our study, we observed that the length of hospital stays of patients hospitalized due to mushroom poisoning and who underwent hemoperfusion with a carbon filter was similar to other studies.

The first hemoperfusion treatment was performed in 1948 by Muirhead and Reid to remove uremic toxins from the circulation. It has been used in drug poisoning since 1951 (9). Although there has been no study since then that has revealed the real place of extracorporeal treatment in mushroom poisoning cases, the United States Association of Poison Control Centers reported that extracorporeal treatment was applied in 0.15-0.22% of all poisoning cases in their data covering the years 1985-1994 (9). Mydlík et al. (15) recommended the use of hemoperfusion filters containing Amberlite® XAD-2 within the first 24-36 hours of acute poisoning with Amanita phalloides. Monhart (16) stated that hemoperfusion reduces hepatic and renal damage, lowers the risk of mortality, and the use of Amberlite® XAD-2 instead of activated charcoal increases the chance of success in treatment. Splendiani et al. (17) applied activated charcoal hemoperfusion to one of two patients with Amanita phalloides poisoning for 3 hours every day for 5 days; to the other patient, continuous venovenous hemodialysis for 20 hours every day for 3 days, and discharged both patients on the 10th day. In our study, 45.1% of the patients started hemoperfusion with a carbon filter within the first 0-20 hours. In 54.9% of the patients, hemodialysis with a carbon filter was started after 21 hours. The mechanism by which hemoperfusion exerts its detoxifying effects includes the removal of not only amatoxins but also inflammatory mediators and hepatotoxic metabolites. Monhart (16) emphasized its potential in reducing hepatic injury when applied within the first 24 hours post-ingestion.

Hemoperfusion, which we use in the treatment of our patients, is done by passing the blood through a carbon-

containing filter. Agents that are highly protein-bound and lipid-soluble are easier to remove from the circulation by hemoperfusion (9). The beneficial effect of hemoperfusion is not only the clearance of alpha-amanitin from plasma, but also the clearance of neurotoxic substances such as methionine, tryptophan, and phenylalanine. Hepatic encephalopathy regresses in 75% of patients with hemoperfusion (1). In the study conducted by Aji et al. (18) in our country, it was reported that thrombocytopenia did not develop in patients after hemoperfusion, but melena developed for 2-3 days. Sabeel et al. (19) reported in their study on 41 patients that the platelet count decreased by 24% after hemodialysis and hemoperfusion were combined, the lowest platelet count was 26,000 after hemodialysis and hemoperfusion were completed, and the decrease in platelets due to hemoperfusion treatment was temporary, thrombocytopenia usually resolved within 24-48 hours and was very rarely associated with bleeding. Hemoperfusion-related thrombocytopenia is generally attributed to platelet adsorption by the activated charcoal filter. This phenomenon, while transient, necessitates close monitoring, particularly in patients with baseline coagulopathy or gastrointestinal bleeding risk (18,19). In our study, it was observed that platelet values decreased after the first day after hemoperfusion with a carbon filter in patients, and the average of the values with the lowest platelet value was 58.51 10^3/μL and the average of the days when this value was seen was 2.5 days. It was observed that the patients' platelet values returned to normal in an average of 6.62 days and the average platelet count on these days was 198.7 10^3/μL. In the present study, like other studies, we observed that platelet values decreased after hemoperfusion with a carbon filter and platelet values returned to normal within days without causing any complications.

In the study conducted by Aji et al. (18), it was reported that the highest liver enzyme levels after hemoperfusion were seen on the second and third days of treatment, followed by a gradual decrease. In our study, the highest value of all biochemical values was observed before or on the first day after hemoperfusion with a carbon filter. A decrease in values was observed within days after hemoperfusion with a carbon filter. In our study, like other studies, we observed improvement in biochemical and hematological values after the application of carbon filter hemoperfusion in patients with mushroom poisoning.

Study Limitations

The primary limitation of our study is the absence of a control group. The lack of a comparator arm restricts the ability to establish causality. However, the standard clinical practice and ethical mandates at our institution require early hemoperfusion for all patients with suspected amatoxin-related mushroom poisoning. Therefore, randomization or withholding hemoperfusion was not feasible. Future prospective studies with matched cohorts are needed to delineate its independent effect.

Conclusion

In our study, hematological and biochemical parameters were evaluated after carbon filter hemoperfusion in patients with mushroom poisoning. We found that there was a statistically significant decrease in the hematological and biochemical parameters evaluated in our study after hemoperfusion with a carbon filter. When hematological parameters were examined after hemoperfusion with a carbon filter in patients with mushroom poisoning, a decrease in platelet values was observed in all patients and increased to normal values within days without causing complications. We observed that it is beneficial to apply hemoperfusion with a carbon filter as soon as possible after the onset of symptoms in patients with mushroom poisoning.

Ethics

Ethics Committee Approval: Local Ethics Committee approval of University of Health Sciences Turkey, Ümraniye Training and Research Hospital (no: B.10.1.TKH.4.34.H.GP.0.01/363, date: 16.12.2021) was obtained.

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

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.B.D., S.B., Concept: M.B.D., S.B., Design: M.B.D., S.B., R.S., Data Collection or Processing: M.B.D., Analysis or Interpretation: M.B.D., S.M.T., R.S., Literature Search: M.B.D., S.M.T., R.S., S.B., Writing: M.B.D., S.M.T., R.S.

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

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

References

1. Balice G, Boksebeld M, Barrier Q, Boccalini S, Kassai-Koupai B, Paret N, et al. Mushroom poisoning-related cardiac toxicity: a case report and systematic review. Toxins (Basel). 2024;16(6):265.

2. Gouvinhas I, Silva J, Alves MJ, Garcia J. The most dreadful mushroom toxins: a review of their toxicological mechanisms, chemical structural characteristics, and treatment. EXCLI J. 2024;23:833-859.

3. Kayes T, Ho V. *Amanita phalloides*-associated liver failure: molecular mechanisms and management. Int J Mol Sci. 2024;25(23):13028.

4. Deniz T, Kandiş H, Saygun M, Büyükkoçak Ü, Ülger H, Karakuş A. Evaluation of intoxication cases applied to emergency department of Kirikkale University Hospital. Duzce Medical Journal. 2009;11(2):15-20.

5. Akkuzu MZ, Sezgin O, Yaraş S, Özdoğan O, Üçbilek E, Altıntaş E, et al. Wild mushroom poisoning in Mersin: a retrospective epidemiological study and investigation of liver toxicity. Van Sag Bil Derg. 2020;13(3):10-14.

6. Nenov VD, Marinov P, Sabeva J, Nenov DS. Current applications of plasmapheresis in clinical toxicology. Nephrol Dial Transplant. 2003;18(Suppl 5):v56-v58.

7. Mydlík M, Derzsiová K, Frank K. Renal replacement therapy in acute poisonings--one center experience. Przegl Lek. 2013;70(6):381-385.

8. Ke J, Wei Y, Chen B. Application of hemoperfusion in the treatment of acute poisoning. Blood Purif. 2024;53(1):49-60.

9. Zhang H, Jiang F, Tang C, Liu Y, Zhang J. Prospects and applications of efficient physical field processing technologies for polysaccharide extraction and quality improvement in edible mushrooms: a systematic review. Int J Biol Macromol. 2025;301:140412.

10. Turan Gökçe D, Arı D, Ata N, Gökcan H, İdilman R, Ülgü MM, et al. Mushroom intoxication in Türkiye: a nationwide cohort study based on demographic trends, seasonal variations, and the impact of climate change on incidence. Turk J Gastroenterol. 2024;36(1):61-66.

11. Iskander P, Marzouk D, Iskander A, Oza F. Malicious mushrooms. Gastro Hep Adv. 2023;2(4):544-546.

12. Tuğcan MO, Akpınar AA. Mushroom poisoning: an updated review. Turk J Emerg Med. 2025;25(1):10-16.

13. Hocaoğlu N, Kalkan Ş, Tunçok Y. Mushroom poisoning reported to the Dokuz Eylül University drug and poison information center. Turk J Emerg Med. 2010;10(3):119-125.

14. Colak S, Kandis H, Afacan MA, Erdogan MO, Gunes H, Kaya E, et al. Assessment of patients who presented to the emergency department with mushroom poisoning. Hum Exp Toxicol. 2015;34(7):725-731.

15. Mydlík M, Derzsiová K, Mizla P, Beno P. Pouzitie hemoperfúzie pri otrave hubami. Klinickýrozbor 58 chorých [Hemoperfusion in mushroom poisoning. Clinical analysis of 58 patients]. Cas Lek Cesk. 1993;132(15):464-467. Slovak.

16. Monhart V. Otravy muchomůrkami a význam sorpcní hemoperfuze pri jejich lécení [Amanita poisoning and the importance of sorption hemoperfusion in its therapy]. Vnitr Lek. 1997;43(10):686-690. Czech.

17. Splendiani G, Zazzaro D, Di Pietrantonio P, Delfino L. Continuous renal replacement therapy and charcoal plasmaperfusion in treatment of amanita mushroom poisoning. Artif Organs. 2000;24(4):305-308.

18. Aji DY, Calişkan S, Nayir A, Mat A, Can B, Yaşar Z, et al. Haemoperfusion in Amanita phalloides poisoning. J Trop Pediatr. 1995;41(6):371-374.

19. Sabeel AI, Kurkus J, Lindholm T. Intensive hemodialysis and hemoperfusion treatment of Amanita mushroom poisoning. Mycopathologia. 1995;131(2):107-114.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):198-205
DOI: 10.4274/BMB.galenos.2026.69775

Comparison of Transobturator Tape (TOT) and BURCH Colposuspension in the Surgical Treatment of Female Urinary Incontinence

Kadınlarda İdrar Kaçırmanın Cerrahi Tedavisinde Transobturator Bant (TOT) ve BURCH Kolposüspansiyonunun Karşılaştırılması

Çağcıl Yetim[1], Ali Ulvi Hakverdi[2]

[1]Medicana Ataköy Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
[2]İzmir Tınaztepe University Galen Hospital, Department of Obstetrics and Gynecology, İzmir, Turkey

Abstract

Objective: The aim of this prospective comparative study is to compare the outcomes of the transobturator tape (TOT) and BURCH operations through clinical and quality of life analyses in women with urinary incontinence.

Method: The patients who presented with urinary incontinence symptoms to the obstetrics and gynecology clinic of a university hospital between June 2007 and June 2008 were included in the study. Patients diagnosed with stress urinary incontinence (SUI) or mixed urinary incontinence (MUI) were scheduled for TOT or BURCH procedures. Preoperative urological and gynecological histories were obtained, along with gynecological examinations, ultrasonography and quality of life questionnaires. All patients were followed for 18 months in the postoperative period. Follow-up studies included stress tests, presence of incontinence and quality of life questionnaires, and postvoid residual volume measurements. Two treatment procedures and their outcomes were compared.

Results: Of the 50 patients, 37 (74%) had SUI, and 13 (26%) had MUI. The mean age of the patients was 47.42±9.48 years (similar within groups). Postoperative decreases in hemoglobin concentration were similar between groups. Postoperative complications occurred in 12 (48%) patients in the TOT group and 5 (20%) patients in the BURCH group (p=0.072). All quality of life scores showed significant improvement between pre-postoperative periods, with similar changes in both groups. Subjective improvement was observed in 88% of patients in the TOT group and 80% in the BURCH group.

Öz

Amaç: Bu prospektif karşılaştırmalı çalışmanın amacı, idrar kaçırma sorunu olan kadınlarda transobturator bant (TOT) ve BURCH operasyonlarının sonuçlarını klinik ve yaşam kalitesi analizleri yoluyla karşılaştırmaktır.

Yöntem: Haziran 2007 ile Haziran 2008 arasında bir üniversite hastanesinin kadın hastalıkları ve doğum kliniğine idrar kaçırma semptomlarıyla başvuran hastalar çalışmaya dahil edildi. Stres üriner inkontinans (SUI) veya karma üriner inkontinans (MUI) tanısı konulan hastalara TOT veya BURCH prosedürleri planlandı. Ameliyat öncesi ürolojik ve jinekolojik öyküler, jinekolojik muayeneler, ultrasonografi ve yaşam kalitesi anketleri alındı. Tüm hastalar ameliyat sonrası dönemde 18 ay boyunca takip edildi. Takip çalışmaları stres testleri, inkontinans varlığı ve yaşam kalitesi anketleri ve işeme sonrası kalan hacim ölçümlerini içeriyordu. İki tedavi prosedürü ve sonuçları karşılaştırıldı.

Bulgular: Elli hastanın 37'sinde (%74) SUI, 13'ünde (%26) MUI vardı. Hastaların ortalama yaşı 47,42±9,48 yıldı (gruplar içinde benzer). Ameliyat sonrası hemoglobin konsantrasyonundaki düşüşler gruplar arasında benzerdi. Ameliyat sonrası komplikasyonlar TOT grubunda 12 (%48) hastada ve BURCH grubunda 5 (%20) hastada meydana geldi (p=0,072). Tüm yaşam kalitesi skorları ameliyat öncesi ve sonrası dönemler arasında anlamlı iyileşme gösterdi ve her iki grupta da benzer değişiklikler oldu. TOT grubunda hastaların %88'inde ve BURCH grubunda hastaların %80'inde öznel iyileşme görüldü. TOT grubunda hastaların %84'ünde ve BURCH grubunda hastaların

Address for Correspondence: Çağcıl Yetim, MD, Medicana Ataköy Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey
E-mail: cagcil@yahoo.com **ORCID:** orcid.org/0000-0001-5859-5391
Received: 10.03.2025 **Accepted:** 04.05.2026 **Epub:** 11.05.2026 **Publication Date:** 24.06.2026

Cite this article as: Yetim Ç, Hakverdi AU. Comparison of transobturator tape (TOT) and BURCH colposuspension in the surgical treatment of female urinary incontinence. Bagcilar Med Bull. 2026;11(2):198-205

Abstract

Objective improvement was noted in 84% of the TOT group and 80% of the BURCH group. These results were statistically similar between the groups.

Conclusion: The success rate of BURCH is similar to the TOT procedure. TOT procedure is also a valuable alternative surgical method with a lower complication rate than the BURCH.

Keywords: BURCH, transobturator tape, urinary incontinence, women

Öz

%80'inde nesnel iyileşme kaydedildi. Bu sonuçlar gruplar arasında istatistiksel olarak benzerdi.

Sonuç: BURCH'nin başarı oranı TOT prosedürüne benzerdir. TOT prosedürü ayrıca BURCH'den daha düşük komplikasyon oranına sahip değerli bir alternatif cerrahi yöntemdir.

Anahtar kelimeler: BURCH, kadınlar, idrar kaçırma, transobturator bant

Introduction

Urinary incontinence (UI), defined by the International Continence Society, is the urine leaks without control that causes social or hygienic problems and can be objectively demonstrated (1,2). Stress urinary incontinence (SUI) is the most common subtype, occurring during physical activities that increase intra-abdominal pressure, such as coughing or lifting. It particularly affects many women in the postmenopausal period; while it is seen at a rate of 35% in Europe, it varies between 16.1% and 25.1% in Turkey, depending on age (3,4).

While conservative therapies—including lifestyle modifications, pelvic floor muscle training, and pharmacological interventions—are recommended as first-line management, a significant proportion of women experience persistent symptoms. In such cases, surgical intervention remains the standard of care (5). For decades, BURCH procedure was considered the gold standard surgical treatment for SUI (6). However, the introduction of minimally invasive mid-urethral sling procedures revolutionized surgical practice. Among these, the transobturator tape (TOT) technique, described by Delorme in 2001, has gained widespread acceptance due to its comparable success rates, shorter operative time, and lower risk of perioperative complications (7,8).

The primary aim of this study is to compare the effectiveness, complication rates, and quality-of-life outcomes between BURCH and TOT procedures, thereby providing evidence to guide clinical decision-making in SUI management. Importantly, unlike many prior studies limited by short-term follow-up or a predominant focus on objective cure rates, the present study emphasizes long-term outcomes and incorporates patient-reported quality-of-life measures. In doing so, it addresses a critical gap in the literature and contributes to a more comprehensive understanding of treatment effectiveness from both clinical and patient-centered perspectives.

Materials and Methods

This prospective study was conducted with patients who consecutively presented with UI symptoms to the Department of Obstetrics and Gynecology at Hatay Mustafa Kemal University Hospital between June 2007 and June 2008. The study included women aged ≥18 who agreed to participate in this study. Patients with predominant urge incontinence, diabetes, or neurological deficits were excluded. Following a preliminary evaluation of patients who consecutively presented to our clinic with UI symptoms, decisions were made regarding the appropriate surgical procedures. BURCH procedure were performed in patients who required open abdominal-pelvic surgery for other indications and patients who had not previously undergone open abdominal-pelvic surgery. The TOT procedure was preferred for patients who had previously undergone open abdominal-pelvic surgery. Additionally, patients' preferences for surgical methods were also taken into consideration. A total of 50 patients diagnosed with SUI or mixed urinary incontinence (MUI) were scheduled for either TOT or BURCH procedures. Each surgical procedure group consisted of 25 patients. This study was not randomized, and allocation to surgical procedures was based on clinical indications, prior surgical history, and patient preference. The effectiveness, feasibility, complication rates, and impact on quality of life were descriptively compared between the two groups; however, due to the non-randomized design, these comparisons should not be interpreted as causal.

In a chi-square model with a Type 1 error (alpha) of 0.05, a power ratio (1-beta) of 0.85, and an effect size (w) of 0.5 and degrees of freedom of 2, we predicted that a total of 44 cases would be included in the study. Because of the potential for data loss, we included a total of 50 patients who met the inclusion criteria and agreed to participate in the study. Informed consent was obtained from all participants. The study was approved by the Ethics Committee of Hatay Mustafa Kemal University Tayfur Ata Sökmen Faculty of

Medicine (approval number: 06/04, date: 25.06.2009). This approval was obtained for the analysis of prospectively collected institutional data.

Data Collection and Preorerative Assessment

Demographic characteristics, incontinence questionnaires, quality of life scores [incontinence quality of life scale (I-QOL), incontinence and prolapse symptom inventory-quality of life scale (PSI-QOLS)], clinical findings, neurological assessments (clitoral reflex, anal reflex, cough reflex), stress tests, Q-tip tests, postvoid residual volumes, hospital stay durations, catheterization times, ultrasonography results, materials used (tape, etc.), operation times, and outcomes were recorded.

I-QOL was developed to determine the QOL of patients with UI. The questionnaire items are in the five-category Likert type. The calculated total score is converted to scores between 0 and 100 and consists of three subdimensions (limiting behaviors, psychological impact, limiting social life). Higher scores indicate a better QOL (9). The Turkish validity and reliability study was conducted by Eyigor et al. (10). We used the PSI-QOL, which contains 11 items for prolapse symptoms and 4 items for life domains of social, physical and entertainment activities, and sexual function (11). The Turkish validation and reliability study of PSI-QOL was conducted by Cam et al. (12). Percentage changes in quality-of-life scores were calculated using the formula: (postoperative score-preoperative score)/preoperative score ×100. High percentage changes may occur when baseline scores are low and should therefore be interpreted with caution.

Pelvic examination was performed using the Baden-Walker classification system. Preoperative and postoperative hemoglobin levels, urine cultures, and urinalyses were obtained. SUI diagnosis was confirmed by observing urine leakage during Valsalva maneuver and coughing after bladder filling with 300 mL saline. Bladder neck mobility was considered positive when the Q-tip test demonstrated an angle change greater than 30°. Patients with urinary tract infection were treated and re-evaluated prior to surgery.

Follow-up Schedule

Postoperative evaluations were performed at 1 week, 1 month, 3 months, 6 months, 12 months, and annually thereafter. Follow-up assessments included stress tests, symptom evaluation, quality of life questionnaires, postvoid residual volume, and complication monitoring. Mean follow-up duration was 15.8±5.9 months (range, 6-24 months).

Surgical Procedures

All patients received antibacterial prophylaxis (cefazolin 1 g intravenously preoperatively and postoperatively) and a vaginal suppository (Biokadin®) one day before surgery.

• **TOT procedure:** After bladder emptying, the TOT technique was performed using I-STOP® polypropylene tape (CL Medical). An 18F Foley catheter was inserted at the end of surgery and removed 36 hours later. Spinal anesthesia was used in 32% of patients and general anesthesia in 68%.

• **BURCH procedure:** Performed according to the Petri technique, using non-absorbable sutures (size 0 or 1). All patients received general anesthesia. A 16F Foley catheter was removed 24 hours postoperatively.

Perioperative data included operative time, hospital stay, catheterization duration, and postoperative residual urine volume. Patients were discharged when postvoid residual volume was <100 mL.

Outcomes and Complications

Surgical success was defined as a negative stress test, postvoid residual volume <100 mL, and complete continence. Failure was defined as persistent leakage or positive stress test. Complications within 15 days were classified as early postoperative complications, and those occurring after 15 days as late complications

Statistical Analysis

Statistical analysis was performed using SPSS for Windows 10.0. Descriptive statistics (mean, standard deviation, frequency) were used. Bonferroni-corrected Friedman and Mann-Whitney U tests were used to compare preoperative and postoperative changes between groups. Wilcoxon tests were applied to compare absolute values before and after surgery. Pearson chi-square and Fisher's exact tests were used for categorical data. Statistical significance was set at $p<0.05$ with a 95% confidence interval.

Results

A total of 50 patients were included in the study. Of these, 37 (74%) were diagnosed with SUI, while 13 (26%) had MUI. Among the SUI patients, 22 underwent TOT surgery, and 18 underwent BURCH procedures. For MUI patients, 3 underwent TOT and 7 underwent BURCH procedures. General anesthesia was used in 17 (68%) TOT cases, while spinal anesthesia was applied in 8 (32%). All BURCH procedures were performed under general anesthesia.

Patient Characteristics

The patients' ages ranged from 27 to 70 years, with a mean age of 47.42±9.48. In the TOT group, the mean age was 49.4±10.73, while in the BURCH group, it was 45.44±7.77, with no statistically significant difference (p=0.177). The body mass index (BMI) of all patients ranged from 20.80 to 39.80 (mean 28.66±4.46), with similar mean BMI values in the TOT (29.73±4.54) and BURCH (27.58±4.20) groups (p=0.071).

Table 1 provides an overview of the demographic and clinical characteristics of TOT and BURCH patients. Eighteen patients (36%) were postmenopausal, while 32 (64%) were in their reproductive period. Parity ranged from 2 to 12, with a mean of 4.82±2.11 across all patients. Parity averages were comparable between the TOT (4.76±2.04) and BURCH (4.88±2.22) groups (p=0.953).

Preoperative Q-tip test positivity (>30 degrees) was observed in 20 patients (40%), with 52% positivity in the TOT group and 28% in the BURCH group. The difference in Q-tip positivity was not statistically significant (p=0.086).

Pelvic Organ Prolapse and Surgical History

All patients had some degree of pelvic organ prolapse. Uterine descent was Grade I in 38 (76%) and Grade II in 12 (24%) patients. Figure 1 illustrates the distribution of cystocele, rectocele, and uterine descent grades across the groups. Cystocele grades were distributed as Grade I in 21 (42%), Grade II in 24 (48%), and Grade III in 5 (10%) cases. Rectocele grades were Grade I in 29 (58%), Grade II in 19 (38%), and Grade III in 2 (4%) patients.

Sixteen patients (32%) had a history of gynecological surgeries, including one incontinence surgery. Figure 2 highlights the distribution of prior surgeries by groups. Table 2 shows the distribution of cystocele, rectocele, and uterine descent grades. Additional gynecological surgeries were performed concurrently in 38 patients (76%), with 68% in the TOT group and 84% in the BURCH group.

Operative and Postoperative Findings

The mean operative time was significantly shorter for TOT (20.03±2.83 minutes) compared to BURCH procedures (36.66±9.24 minutes, p<0.001). The mean catheterization time was longer in the TOT group (40.64±21.63 hours)

Table 1. Characteristics of patients in TOT and BURCH groups

Characteristic	All patients (n=50)	TOT group (n=25)	BURCH group (n=25)	p-value
	Mean	Mean	Mean	
Age (years)	47.42	49.40	45.44	0.177
Parity	4.82	4.76	4.88	0.953
BMI (kg/m²)	28.66	29.73	27.58	0.071

TOT: Transobturator tape, BMI: Body mass index

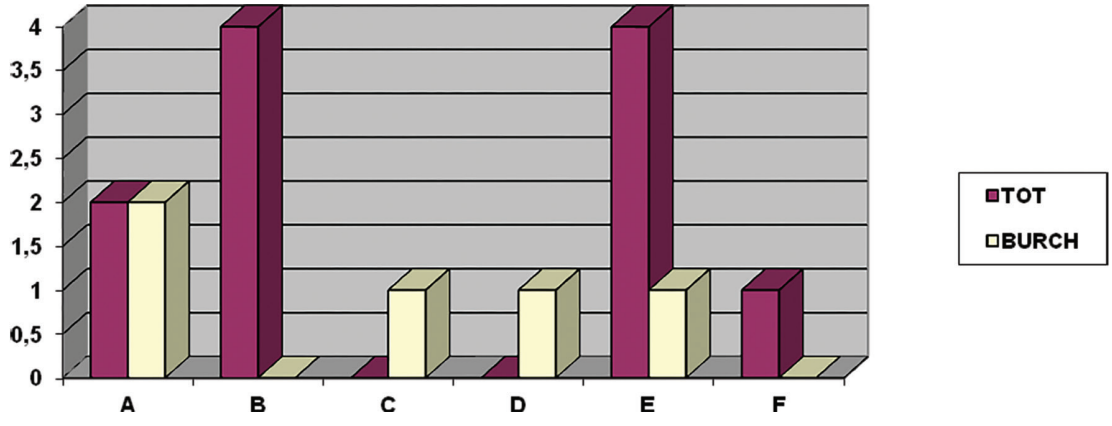


Figure 1. Distribution of prior surgeries by groups

A: Total abdominal hysterectomy with bilateral salpingo-oophorectomy (TAH+BSO); B: Cesarean section; C: Tubal ligation; D: Oophorectomy; E: Colporrhaphy; F: BURCH colposuspension

TOT: Transobturator tape

compared to the BURCH group (28.44±14.28 hours, p<0.001). Longer catheterization time in the TOT group compared to the BURCH group was observed; however, this finding should be interpreted cautiously and does not necessarily indicate a clinically meaningful disadvantage. The mean hospital stay was also shorter for TOT (2.23±1.01 days) than BURCH (2.54±0.59 days, p<0.001).

There was no significant difference in hemoglobin decline postoperatively between the TOT (14.61±8.33%) and BURCH (16.15±6.73%) groups (p=0.383). No major intraoperative complications such as vaginal perforation or nerve injury occurred in TOT cases. However, three BURCH cases experienced intraoperative bleeding managed by suture placement and compression. Table 3 shows the distribution of postoperative complications in TOT and BURCH groups.

Postoperative Complications

Complications included urinary retention in one patient from each group, urinary tract infections in 5 patients (3 TOT, 2 BURCH), and transient groin pain in 5 patients (4 TOT, 1 BURCH), all managed successfully.

Objective and Subjective Outcomes

Objective success was defined as negative stress tests, postvoid residuals <100 mL, and full continence at 3 and 12 months postoperatively. Success rates were similar between the groups (TOT 84% vs. BURCH 80%, p>0.05). Subjective success, assessed using patient-reported improvement scales, was also comparable (TOT 88% vs. BURCH 80%, p>0.05).

Quality of Life

Quality of life improvements were assessed using the I-QOL and PSI-QOLS. Significant postoperative improvements were noted in both scales for both groups (p<0.01). The percentage changes in I-QOL scores from baseline to 12 months were 101.54% for TOT and 57.86% for BURCH. PSI-QOLS scores showed similar trends, with mean improvements of 26.33% for TOT and 31.56% for BURCH, with no significant differences between the groups (p=0.844). Table 4 summarizes the changes in I-QOL and PSI-QOLS survey results.

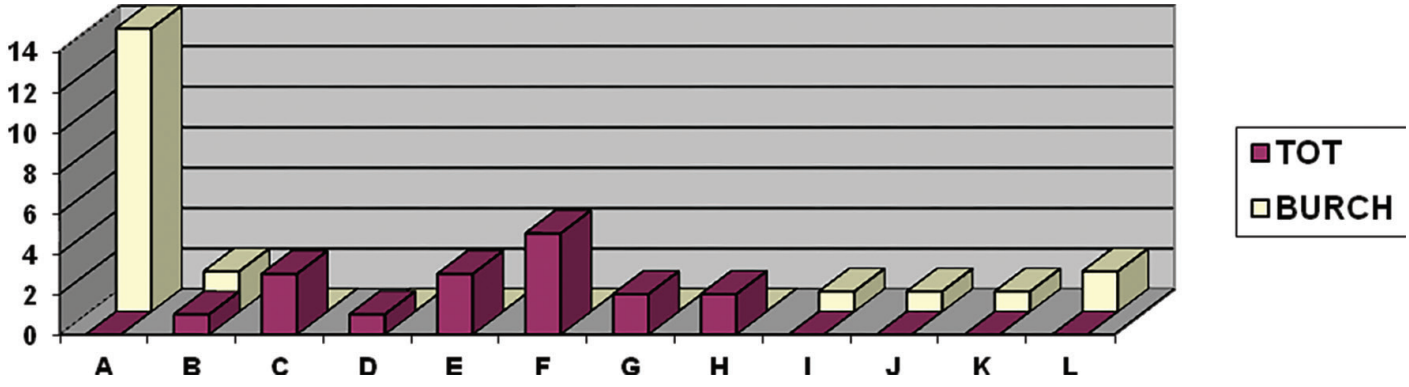


Figure 2. Additional gynecological procedures performed concurrently with transobturator tape (TOT) and BURCH surgeries

***A**: Total abdominal hysterectomy with bilateral salpingo-oophorectomy and culdeplasty (TAH+BSO+culdeplasty); **B**: Total abdominal hysterectomy with bilateral salpingo-oophorectomy (TAH+BSO); **C**: Vaginal hysterectomy with colporrhaphy (VH+CAP); **D**: Colporrhaphy (CAP); **E**: Anterior colporrhaphy (CA); **F**: Posterior colporrhaphy (CP); **G**: Laparoscopic tubal ligation with colporrhaphy (L/S T/L+CAP); **H**: Colpotomy with tubal ligation and colporrhaphy (colpotomy T/L+CP); **I**: Myomectomy with tubal ligation and colporrhaphy (myomectomy T/L+CP); **J**: Oophorectomy; **K**: Mini-laparotomy with tubal ligation (mini laparotomy T/L); **L**: Sacrocolpopexy with culdeplasty*

Table 2. Distribution of cystocele, rectocele, and uterine descent grades

	Grade I n (%)	Grade II n (%)	Grade III n (%)
Cystocele	21 (42%)	24 (48%)	5 (10%)
Rectocele	29 (58%)	19 (38%)	2 (4%)
Uterine descent	38 (76%)	12 (24%)	-

Table 3. Distribution of postoperative complications

	TOT		BURCH	
	Observed n (%)	Not observed n (%)	Observed n (%)	Not observed n (%)
Urinary tract infection	3 (12%)	22 (88%)	2 (8%)	23 (92%)
Urinary retention	1 (4%)	24 (96%)	1 (4%)	24 (96%)
Vaginitis	2 (8%)	23 (92%)	-	25 (100%)
Groin pain	4 (16%)	21 (84%)	1 (4%)	24 (96%)
Dyspareunia	2 (8%)	23 (92%)	-	25 (100%)
***De novo* urge symptoms**	-	25 (100%)	1 (4%)	24 (96%)
Fever	-	25 (100%)	-	25 (100%)
Erosion	-	25 (100%)	-	25 (100%)

TOT: Transobturator tape

Table 4. Changes in I-QOL and PSI-QOLS questionnaire results

	TOT		p**	BURCH		p**
	Preoperative Mean ± SD	Postoperative 12th month Mean ± SD		Preoperative Mean ± SD	Postoperative 12th month Mean ± SD	
*I-QOL	54.26±24.21	92.15±19.10	0.001	58.60±20.24	80.95±19.02	0.001
*PSI-QOLS	50.94±9.77	62.26±7.70	0.001	51.15±13.84	61.95±8.87	0.001

TOT: Transobturator tape, I-QOL: Incontinence quality of life scale, PSI-QOLS: Prolapse symptom inventory-quality of life scale, SD: Standard deviation, *: Friedman test, **: p<0.05

Discussion

UI is a big public health issue with increasing awareness and aging population driving the demand for treatment. Despite many surgical techniques developed, there is no universally accepted gold standard exists for all cases. Among the various types of UI, SUI is the most common and has the most options. But success of these interventions depends on accurate diagnosis of the underlying pathophysiology and choosing the right surgical technique. Our study compared TOT and BURCH procedure, two common surgical methods for SUI, on their effectiveness, safety and patient reported outcomes. TOT appeared to be more frequently preferred in patients with urethral hypermobility, while BURCH was more commonly used in patients requiring concurrent pelvic surgery. This is in line with previous studies that TOT is good for urethral hypermobility and BURCH for complex pelvic floor dysfunction (13,14).

Similar to the literature, TOT had significantly shorter operative time and hospital stay compared to BURCH procedure (15-17). However, unlikely, we found a longer postoperative catheterization time in the TOT group than the BURCH group. Sharma et al. (16) presented a mean of 1.2 day, Roumeguère et al. (17) presented a mean of 0.9 day for postoperative catheterization time in TOT patients. We think that our conflictive result about this issue is likely due to transient voiding dysfunction associated with sling positioning.

In this study, the overall complication rate was similar between the two groups, TOT had slightly higher postoperative complication rate (48% vs. 20%). Although the complication rate was numerically higher in the TOT group, this difference did not reach statistical significance and therefore cannot be interpreted as a meaningful clinical difference. Most of the complications like UTI and transient groin pain were minor and managed conservatively. No major intraoperative complications like bladder perforation or nerve injury occurred in either group. This is in line with previous studies showing the safety of both procedures (18,19).

Objective success defined as negative stress test and postvoid residual under 100 mL, was comparable between the TOT (84%) and BURCH (80%) groups, with no statistically significant difference. Similarly, subjective success, assessed through patient-reported improvement scales, showed nearly equivalent outcomes (88% for TOT vs. 80% for BURCH). These findings reinforce that both surgical techniques yield robust and consistent efficacy in appropriately selected patients, in line with prior reports

demonstrating that the choice of procedure does not substantially alter cure rates when clinical indications are carefully considered (7,20).

Quality-of-life outcomes followed a similar pattern. Both groups experienced significant improvements in I-QOL and PSI-QOL scores postoperatively, reflecting meaningful patient-perceived benefits. Although TOT demonstrated a numerically higher percentage improvement in I-QOL scores, this may be influenced by lower baseline values, and no statistically significant difference was observed between the groups. This indicates that while minor variations in effect size may exist, both procedures contribute meaningfully to patient-reported quality of life improvements (21,22). The lack of significant divergence across both objective and subjective measures highlights the clinical equivalence of TOT and BURCH in terms of overall treatment success.

TOT and BURCH procedures have unique benefits that make them suitable for different patient groups. TOT is minimally invasive, has a shorter surgery time, and can be done under regional anesthesia, making it a good option for older and obese patients. On the other hand, BURCH procedure is better suited for patients who need additional pelvic procedures or have anatomical issues that make TOT unsuitable. These observations are consistent with studies that highlight the need for personalized treatment plans based on individual patient characteristics and clinical situations (14,23). Although TOT had higher rates of postoperative complications, most were minor and temporary. Common issues included urinary retention and groin pain, both of which resolved with conservative treatment. BURCH procedure had a lower complication rate but showed a slightly higher occurrence of urinary retention that required catheterization. Both techniques had low rates of severe complications, such as urethral or bladder injuries, which aligns with established safety profiles (23).

Recent long-term studies indicate that while TOT shows better short-term outcomes in terms of surgery time and reduced hospital stays, the modified laparoscopic BURCH procedure provides durable benefits for selected patient groups with specific anatomical considerations. Our findings support these observations, showing comparable objective success rates between both methods, which highlights the need for tailoring surgical choice to individual patient profiles rather than adopting a universal approach (24).

Clinical considerations also point to the importance of minimizing complications such as mesh erosion and chronic pain. Although TOT was associated with more transient minor complications, the modified BURCH approach demonstrated lower rates of dyspareunia and chronic pelvic pain, consistent with prior reports (25). These patterns reinforce that both procedures remain viable options, with distinct strengths that should be matched to patient-specific factors.

Study Limitations

This study has several limitations. First, the study design was non-randomized, and allocation to surgical procedures was influenced by clinical indications, prior surgical history, and patient preference. Therefore, selection bias cannot be excluded, and the two groups may not be fully comparable. Accordingly, direct comparative or causal interpretations between TOT and BURCH procedures should be made with caution. Furthermore, the relatively small sample size limits the statistical power of the study and the generalizability of the findings.

Conclusion

In conclusion, both TOT and BURCH procedure are effective and safe surgical options for managing SUI, with similar objective and subjective success rates. TOT may be considered a suitable option for patients seeking a minimally invasive procedure with faster recovery, while BURCH is more suitable for those requiring concurrent pelvic surgery or at higher risk of mesh-related complications. While both techniques are effective and safe, patient-specific anatomical and clinical features should guide surgical choice. Surgical decision-making should therefore be guided by patient anatomy, comorbidities, and treatment priorities. Future randomized studies with longer-term follow-up may further clarify optimal patient selection criteria.

Ethics

Ethics Committee Approval: The study was approved by the Ethics Committee of Hatay Mustafa Kemal University Tayfur Ata Sökmen Faculty of Medicine (approval number: 06/04, date: 25.06.2009).

Informed Consent: Informed consent was obtained from all participants.

Acknowledgments

We would like to thank our nurses and staff who assisted us during our surgeries.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ç.Y., A.U.H., Concept: Ç.Y., A.U.H., Design: Ç.Y., A.U.H., Data Collection or Processing: Ç.Y., A.U.H., Analysis or Interpretation: Ç.Y., A.U.H., Literature Search: Ç.Y., Writing: Ç.Y., A.U.H.

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

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

References

1. Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P, Ulmsten U, et al.; Standardisation Sub-Committee of the International Continence Society. The standardisation of terminology in lower urinary tract function: report from the standardisation sub-committee of the International Continence Society. Urology. 2003;61(1):37-49.
2. Abrams P, Cardozo L, Fall M, Griffiths D, Rosier P, Ulmsten U, et al.; Standardisation Sub-committee of the International Continence Society. The standardisation of terminology of lower urinary tract function: report from the standardisation sub-committee of the International Continence Society. Neurourol Urodyn. 2002;21(2):167-178.
3. Milne JS, Williamson J, Maule MM. Urinary symptoms in older people. Modern Geriatrics. 1972;2:198.
4. Biri A, Durukan E, Maral I, Korucuoğlu U, Biri H, Týraş B, et al. Incidence of stress urinary incontinence among women in Turkey. Int Urogynecol J Pelvic Floor Dysfunct. 2006;17(6):604-610.
5. Nambiar AK, Bosch R, Cruz F, Lemack GE, Thiruchelvam N, Tubaro A, et al. EAU guidelines on assessment and nonsurgical management of urinary incontinence. Eur Urol. 2018;73(4):596-609.
6. Dainer M, Hall CD, Choe J, Bhatia NN. The Burch procedure: a comprehensive review. Obstet Gynecol Surv. 1999;54(1):49-60.
7. Delorme E. La bandelette trans-obturatrice: un procédé mini-invasif pour traiter l'incontinence urinaire d'effort de la femme [Transobturator urethral suspension: mini-invasive procedure in the treatment of stress urinary incontinence in women]. Prog Urol. 2001;11(6):1306-1313. French.
8. Latthe PM, Foon R, Toozs-Hobson P. Transobturator and retropubic tape procedures in stress urinary incontinence: a systematic review and meta-analysis of effectiveness and complications. BJOG. 2007;114(5):522-531. Erratum in: BJOG. 2007;114(10):1311.
9. Patrick DL, Martin ML, Bushnell DM, Yalcin I, Wagner TH, Buesching DP. Quality of life of women with urinary incontinence: further development of the incontinence quality of life instrument (I-QOL). Urology. 1999;53(1):71-76. Erratum in: Urology 1999;53(5):1072.
10. Eyigor S, Karapolat H, Akkoc Y, Yesil H, Ekmekci O. Quality of life in patients with multiple sclerosis and urinary disorders: reliability and validity of Turkish-language version of incontinence quality of life scale. J Rehabil Res Dev. 2010;47(1):67-71.
11. Digesu GA, Khullar V, Cardozo L, Robinson D, Salvatore S. P-QOL: a validated questionnaire to assess the symptoms and quality of life of women with urogenital prolapse. Int Urogynecol J Pelvic Floor Dysfunct. 2005;16(3):176-181; discussion 181.
12. Cam C, Sakalli M, Ay P, Aran T, Cam M, Karateke A. Validation of the prolapse quality of life questionnaire (P-QOL) in a Turkish population. Eur J Obstet Gynecol Reprod Biol. 2007;135(1):132-135.
13. Faerber GJ. Endoscopic collagen injection therapy in elderly women with type I stress urinary incontinence. J Urol. 1996;155(2):512-514.
14. Jeon MJ, Chung da J, Park JH, Kim SK, Kim JW, Bai SW. Surgical therapeutic index of tension-free vaginal tape and transobturator tape for stress urinary incontinence. Gynecol Obstet Invest. 2008;65(1):41-46.
15. Sivaslioglu AA, Caliskan E, Dolen I, Haberal A. A randomized comparison of transobturator tape and Burch colposuspension in the treatment of female stress urinary incontinence. Int Urogynecol J Pelvic Floor Dysfunct. 2007;18(9):1015-1019.
16. Sharma JB, Thariani K, Kumari R, Kaur T, Uppal B, Pandey K, et al. Short and long term follow up and efficacy of trans obturator tape for management of stress urinary incontinence. J Obstet Gynaecol India. 2021;71(3):285-291.
17. Roumeguère T, Quackels T, Bollens R, de Groote A, Zlotta A, Bossche MV, et al. Trans-obturator vaginal tape (TOT) for female stress incontinence: one year follow-up in 120 patients. Eur Urol. 2005;48(5):805-809.
18. Galloway NT, Davies N, Stephenson TP. The complications of colposuspension. Br J Urol. 1987;60(2):122-124.
19. Saidi MH, Gallagher MS, Skop IP, Saidi JA, Sadler RK, Diaz KC. Extraperitoneal laparoscopic colposuspension: short-term cure rate, complications, and duration of hospital stay in comparison with Burch colposuspension. Obstet Gynecol. 1998;92(4 Pt 1):619-621.
20. de Leval J. Novel surgical technique for the treatment of female stress urinary incontinence: transobturator vaginal tape inside-out. Eur Urol. 2003;44(6):724-730.
21. Holmgren C, Nilsson S, Lanner L, Hellberg D. Long-term results with tension-free vaginal tape on mixed and stress urinary incontinence. Obstet Gynecol. 2005;106(1):38-43.
22. Falconer C, Söderberg M, Blomgren B, Ulmsten U. Influence of different sling materials on connective tissue metabolism in stress urinary incontinent women. Int Urogynecol J Pelvic Floor Dysfunct. 2001;12(Suppl 2):S19-S23.
23. Demirci F, Petri E. Perioperative complications of Burch colposuspension. Int Urogynecol J Pelvic Floor Dysfunct. 2000;11(3):170-175.
24. Chen Y, Zhang C, Yang S, Chen J, Peng L, Chen J, et al. Long-term outcomes of surgical interventions for stress urinary incontinence: a systematic review and network meta-analysis. Int J Surg. 2024;110(1):520-528.
25. Brasoveanu S, Balulescu L, Grigoraş D, Erdelean D, Olaru F, Bardan R, et al. Evaluating patient preferences and clinical outcomes for modified laparoscopic Burch colposuspension and transobturator tape procedures in stress urinary incontinence treatment. Life (Basel). 2024;14(3):380.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):206-212
DOI: 10.4274/BMB.galenos.2026.28291

Assessing Radial Nerve Entrapment and Traumatic Radial Neuropathy in the Arm Segment: Clinical Correlations and Electrodiagnostic Findings

Üst Kol Segmentinde Radyal Sinir Sıkışması ve Travmatik Radyal Nöropatinin Değerlendirilmesi: Klinik Korelasyonlar ve Elektrodiagnostik Bulgular

Şencan Buturak[1], Halit Fidancı[2], Halil Can Alaydın[2], Zülfikar Arlıer[1]

[1]University of Health Sciences Turkey, Adana City Training and Research Hospital, Department of Neurology, Adana, Turkey
[2]University of Health Sciences Turkey, Adana City Training and Research Hospital, Department of Neurology, Division of Clinical Neurophysiology, Adana, Turkey

Abstract

Objective: The entrapment of the radial nerve at the spiral groove of the humerus in the arm segment (ERNAS) and traumatic radial neuropathy in the arm segment (TRNAS) are disorders that present with dropped hand. The aim of this study was to compare electrodiagnostic and clinical findings between patients with ERNAS and TRNAS, and to evaluate whether there is a correlation between these findings.

Method: This study was conducted at the Clinical Neurophysiology Laboratory of University of Health Sciences Turkey, Adana City Training and Research Hospital. ERNAS and TRNAS patients were included. The electrodiagnostic findings and the disabilities of the arm, shoulder and hand questionnaire (DASH) scores of the patients were reviewed in this retrospective study.

Results: Twenty-two ERNAS (20 male, 2 female) and 8 TRNAS (5 male, 3 female) patients were included. The mean ages of ERNAS and TRNAS patients were 45.5±16.1 and 45.0±13.8 years, respectively. The amplitudes of the superficial radial nerve compound nerve action potential (CNAP) and the radial nerve compound muscle action potential (CMAP) were significantly reduced in patients with TRNAS

Öz

Amaç: Üst kol segmentinde humerus spiral oluğunda radyal sinir sıkışması (RSS) ve travmatik radyal nöropati (TRN), düşük el ile seyreden bozukluklardır. Bu çalışmanın amacı, RSS ve TRN hastalarında elektrodiagnostik ve klinik bulguları karşılaştırmak ve bu bulgular arasında bir korelasyon olup olmadığını değerlendirmektir.

Yöntem: Çalışma, Sağlık Bilimleri Üniversitesi, Adana Şehir Eğitim ve Araştırma Hastanesi, Klinik Nörofizyoloji Laboratuvarı'nda gerçekleştirildi. RSS ve TRN tanısı alan hastalar çalışmaya dahil edildi. Retrospektif olarak tasarlanan çalışmada, hastaların elektrodiagnostik bulguları ve kol, omuz ve el sorunları anketi (DASH) skorları değerlendirildi.

Bulgular: Yirmi iki RSS (20 erkek, 2 kadın) ve sekiz TRN (5 erkek, 3 kadın) hastası çalışmaya dahil edildi. RSS ve TRN hastalarının yaş ortalamaları sırasıyla 45,5±16,1 ve 45,0±13,8 yıl idi. TRN hastalarında, yüzeyel radyal sinir bileşik sinir aksiyon potansiyeli (BSAP) ve radyal sinir bileşik kas aksiyon potansiyeli (BKAP) amplitüdleri, RSS hastalarına kıyasla anlamlı düzeyde düşük bulundu (sırasıyla p=0,002 ve p=0,007). RSS hastalarında DASH skorları ile yüzeyel radyal duyu sinir iletim hızı (SİH) arasında negatif korelasyon saptandı (p=0,001, r=-0,677).

Address for Correspondence: Assoc. Prof., Halit Fidancı, University of Health Sciences Turkey, Adana City Training and Research Hospital, Department of Neurology, Division of Clinical Neurophysiology, Adana, Turkey
E-mail: dr.halitfidanci@gmail.com **ORCID:** orcid.org/0000-0001-6573-9090
Received: 26.03.2025 **Accepted:** 02.05.2026 **Epub:** 07.05.2026 **Publication Date:** 24.06.2026

Cite this article as: Buturak Ş, Fidancı H, Alaydın HC, Arlıer Z. Assessing radial nerve entrapment and traumatic radial neuropathy in the arm segment: clinical correlations and electrodiagnostic findings. Bagcilar Med Bull. 2026;11(2):206-212

Abstract

compared to those with ERNAS (p=0.002 and p=0.007, respectively). A negative correlation was identified between DASH scores and superficial radial sensory nerve conduction velocity (NCV) in ERNAS patients (p=0.001, r=-0.677).

Conclusion: This study indicated that superficial radial nerve CNAP and radial nerve CMAP amplitudes were more affected in TRNAS patients than in ERNAS patients. It was also concluded that there may be a relationship between superficial radial sensory NCV and clinical features.

Keywords: Electrodiagnosis, nerve conduction study, radial neuropathy, saturday night palsy, trauma

Öz

Sonuç: Bu çalışma, TRN hastalarında yüzeyel radyal sinir BSAP ve radyal sinir BKAP amplitüdlerinin RSS hastalarına göre daha fazla etkilendiğini ortaya koymuştur. Ayrıca, yüzeyel radyal duyu SİH ile klinik özellikler arasında bir ilişki olabileceği sonucuna varılmıştır.

Anahtar kelimeler: Cumartesi gecesi felci, elektrodiagnozis, radyal nöropati, sinir iletim çalışması, travma

Introduction

Radial nerve injury in the arm segment (RNIAS) may lead to sensory disturbances in the radial nerve distribution or with dropped hand, potentially leading to functional impairment of the upper extremity (1-4). RNIAS may occur due to trauma, such as a humeral fracture, or from compression of the radial nerve at the spiral groove of the humerus (1-4). Electrodiagnostic findings, along with clinical features, are essential for the diagnosis of RNIAS (1,2). Electrodiagnostic tests crucial for diagnosis of RNIAS include conduction block of the radial motor nerve, slowing of radial motor nerve conduction velocity (NCV) across the arm segment, and needle electromyography (EMG) abnormalities in radial nerve innervated muscles (1,2,5). Patients with the entrapment of the radial nerve at the spiral groove of the humerus in the arm segment (ERNAS) and traumatic radial neuropathy in the arm segment (TRNAS) may have similar clinical and electrodiagnostic features. Therefore, this study focused on evaluating the clinical and electrodiagnostic characteristics in patients with TRNAS and ERNAS, as well as investigating the relationship between these features. Consequently, understanding the pathophysiology of RNIAS was also an objective.

Materials and Methods

Patients and Study Design

Patients exhibiting clinical and electrodiagnostic characteristics consistent with RNIAS, who were referred to the Clinical Neurophysiology Laboratory at University of Health Sciences Turkey, Adana City Training and Research Hospital between November 2018 and November 2022, were included. Clinical characteristics, findings of electrodiagnostic tests, and imaging methods of all patients were recorded in this retrospective study. Assessment of muscular function was achieved with the medical research council (MRC) scale (6). The disabilities of the arm, shoulder and hand questionnaire (DASH) were applied to all patients to evaluate upper extremity functions (7,8). Additionally, the DASH work module and sports/performing arts module scores were recorded. The DASH questionnaire consists of 30 items assessing upper extremity function, each rated from 1 (no difficulty) to 5 (unable to perform), with the total score ranging from 0 (no disability) to 100 (severe disability). To obtain a valid score, at least 27 items must be completed (7,8). The optional DASH work and sports/performing arts modules each include four additional items, designed for individuals engaged in those specific activities. Patients were considered to have RNIAS if the following features were present (2-5,9,10): 1) Dropped hand/fingers and/or sensory impairment in the cutaneous region innervated by the radial nerve on neurological examination; and 2) Radial motor NCV slowing or motor conduction block of the radial nerve in the arm segment and/or needle abnormality in radial nerve innervated muscles other than the triceps muscle. Patients with the listed conditions were excluded from the study: 1) Mononeuropathy other than radial neuropathy; 2) History of diseases that can lead to polyneuropathy, including diabetes mellitus; 3) Polyneuropathy; 4) Neurodegenerative disease; and 5) Clinical features, electrodiagnostic results, and imaging methods were compatible with radiculopathy or brachial plexopathy. Patients who met the diagnostic criteria for RNIAS were divided into two groups according to the mechanism of onset. Patients who met the diagnostic criteria for RNIAS were included in the ERNAS group when symptom onset occurred after prolonged external compression during sleep, in the absence of any history

of traumatic injury to the upper extremity. Patients with RNIAS whose symptom onset occurred after a documented traumatic event involving the upper extremity, specifically a humeral fracture, were included in the TRNAS group. Ethics committee approval was received from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Adana City Training and Research Hospital (number: 118/2303, date: 15.12.2022).

Electrodiagnostic Tests

Nerve conduction study and needle EMG were performed on the Cadwell Sierra Summit EMG unit (Cadwell Laboratories, Kennewick, Washington, USA). Electrodiagnostic tests were performed if the extremity temperature was 32 degrees or above. In motor and sensory nerve conduction study, low-high filters were set as 20 Hz-10 kHz and 20 Hz-2 kHz, respectively. Surface electrodes were employed in stimulation. Recording was done with disc electrodes. For the motor-sensory nerve conduction study, sweep speed and sensitivity were set as 5-1 ms/division and 2000-10 µV/division, respectively. Recommended methods were used for median, ulnar and superficial radial nerve conduction studies (11-13). In median and ulnar motor nerve conduction studies, recordings were obtained from the abductor pollicis brevis and abductor digiti minimi muscles, respectively. Median and ulnar sensory nerve conduction studies were performed antidromically in the index finger-wrist and little finger-wrist segments, respectively. Amplitudes were measured peak to peak, and sensory nerve conduction velocities were calculated using peak latency. Normal values from the test laboratory were used for reference values in median and ulnar nerve conduction studies (13). For superficial radial sensory nerve conduction studies, recordings were obtained from the anatomical snuffbox region, with the reference electrode placed on the thumb. The nerve was stimulated 12-14 cm proximal to the recording electrode along the radial aspect of the forearm, and sensory NCV was calculated using the peak latency. The lower reference value for superficial radial nerve NCV and compound nerve action potential (CNAP) amplitude was set as 35.7 m/s, 11 µV, respectively (12). Superficial radial nerve CNAP amplitude was considered abnormal if the superficial radial nerve CNAP amplitude was lower than the reference value or reduced by more than 50% compared to the contralateral superficial radial nerve CNAP amplitude. In the radial motor nerve conduction study, recording was conducted using a concentric electrode (length=50 mm, diameter=0.46 mm, Bionen Medical Devices, Florence, Italy; length=50 mm, 26 G, Natus, Galway, Ireland). The recording was made in the extensor indicis proprius muscle. The stimulation points were 4 cm proximal to the recording electrode, 5-6 cm proximal to the lateral epicondyle and the ERB point. The lower reference values for radial motor NCV for the distal and proximal segments were 49.8 m/s and 59.4 m/s, respectively (11). If radial nerve compound muscle action potential (CMAP) could not be obtained, the radial nerve CMAP amplitude was considered abnormal. Motor conduction block was identified as a reduction of more than 50% in the CMAP amplitude from proximal stimulation relative to the CMAP amplitude from distal stimulation.

Needle EMG was applied with concentric needle electrodes. The low-high filter was set to 10 Hz-10 kHz. Sensitivity and sweep speed for spontaneous activity were 100 µV/division and 10 ms/division, respectively. Sensitivity and sweep speed for motor unit action potential (MUAP) examination were 0.5-1 mV/division and 10 ms/division, respectively. The presence of spontaneous discharges at rest was carefully evaluated. MUAP amplitude >3.5 mV or MUAP duration >15 ms was considered neurogenic MUAP. Needle EMG was applied to the extensor indicis proprius, extensor digitorum communis, brachioradialis, triceps, first dorsal interosseous and deltoid muscles in all patients. Additionally, in some patients, the abductor pollicis brevis, biceps brachii and cervical paraspinal muscles were examined with needle EMG for differential diagnosis.

Statistical Analysis

Nominal variables were described using rates and percentages. Numerical variables were reported as mean ± standard deviation, median, and minimum-maximum. Pearson's chi-square test was used to analyze categorical data between groups. The Mann-Whitney U test was employed for comparing numerical variables across groups. Spearman rank correlation analysis was conducted to assess relationships between variables. A p-value of less than 0.05 was considered statistically significant. Statistical analysis was performed using SPSS version 22.0.

Results

Among the 44 reviewed RNIAS patients (26 ERNAS, 18 TRNAS), 14 were excluded from the study. Two of the ERNAS patients had carpal tunnel syndrome, one had diabetes mellitus, and one had cervical radiculopathy. Three of the TRNAS patients had diabetes mellitus, four had carpal tunnel syndrome, one had ulnar neuropathy, one had cervical radiculopathy, and one had polyneuropathy associated with thyroid disease. These patients were excluded from the study. Ultimately, twenty-five male and five female RNIAS

patients were included. The mean age and body mass index of the patients were 45.4±5.3 (min-max 22-69) years and 24.8±2.9 (18.8-31.2) kg/m^2, respectively. The mean duration of illness was 55.4±41.9 (15-180) days. There were 22 ERNAS (20 male, 2 female) and 8 TRNAS patients (5 male, 3 female) (p=0.101 for gender). The mean ages (min-max) of ERNAS and TRNAS patients were 45.5±16.1 (22-69) and 45.0±13.8 (29-67) years, respectively (p=0.907). The mean durations of disease in TRNAS and ERNAS patients were 103.4±50.4 and 38.1±19.9 days, respectively (p<0.005).

RNIAS was identified in 17 patients affecting the right arm and in 13 patients affecting the left arm. Five ERNAS patients had a history of alcohol use, and two ERNAS patients had a history of narcotic use. Other ERNAS patients had a history of radial neuropathy that developed only after prolonged sleep. In 19 of the patients (63%), sensory impairment in the cutaneous area supplied by the superficial radial nerve was present during the neurological assessment. Sensory abnormalities were present in twelve ERNAS patients (55%) and seven TRNAS patients (88%) (p=0.199). All patients had muscle weakness in at least one of the radial innervation muscles. The mean (min-max) MRC scores of finger extension, wrist extension and brachioradialis muscles were 2.0±1.4 (0-4), 2.3±1.7 (0-5), 3.2±1.5 (0-5), respectively. The mean scores of DASH, DASH-work module and DASH-sports performing arts module were 50.2±28.3 (number=30), 57.9±29.3 (number=29), 50.0±27.7 (number=6), respectively. Superficial radial CNAP amplitude and NCV were 22.7±19.6 µV and 52.9±10.9 m/s, respectively. The radial motor nerve distal CMAP amplitude, the NCV and the percentage of amplitude reduction in the arm segment were 4.8±4.0 mV, 51.7±14.5 m/s and 68.9±30.9%, respectively. The comparison of clinical and electrodiagnostic findings between TRNAS and ERNAS patients is presented in Table 1. Table 2 shows the abnormalities found in the electrodiagnostic tests of the patients.

Figure 1 shows ERNAS and TRNAS patients with superficial radial nerve CNAP and radial nerve CMAP abnormalities. Table 3 shows the correlation between nerve conduction study and DASH scores in patients. There was a significant negative correlation between DASH scores and the superficial radial nerve NCV in ERNAS patients and in the overall RNIAS cohort (p=0.001, r=-0.677 for ERNAS patients; p=0.044, r=-0.406 for all RNIAS patients). Figure 2 shows the correlation between superficial radial NCV and DASH scores in ERNAS patients.

Discussion

In this study, the correlation between clinical features and electrodiagnostic findings in RNIAS patients was investigated. A negative correlation was found between DASH scores and superficial radial nerve NCV in ERNAS

Table 1. The clinical and electrodiagnostic features in ERNAS and TRNAS patients

Clinical and electrodiagnostic features	ERNAS Mean ± SD (median) (n)	TRNAS Mean ± SD (median) (n)	p-value
Clinical features			
MRC score of finger extension	1.95±1.49 (3) (n=22)	2±1.19 (2) (n=8)	0.807
MRC score of wrist extension	2.31±1.96 (3) (n=22)	2.37±0.91 (2) (n=8)	0.810
MRC score of the brachioradialis muscle	3.1±1.62 (4) (n=22)	3.37±0.91 (3) (n=8)	0.789
DASH scores	45.4±27.29 (45.8) (n=22)	63.5±28.35 (71.9) (n=8)	0.127
DASH-work module scores	56.21±31.55 (56) (n=21)	62.37±23.37 (62.5) (n=8)	0.805
DASH-sports performing arts module scores	42.18±31.19 (50) (n=4)	65.6±13.29 (65.6) (n=2)	0.355
Electrodiagnostic features			
Superficial radial nerve (µV) CNAP amplitude	28.55±19.51 (24.65) (n=22)	6.76±7.98 (4.35) (n=8)	**0.001**
Superficial radial nerve (m/s) NCV[a]	52±11.14 (50) (n=21)	57.9±9.4 (55.3) (n=4)	0.194
Radial nerve CMAP amplitude	5.83±3.79 (5.3) (n=22)	1.86±3.27 (0) (n=8)	**0.006**
Radial motor NCV across the arm segment[a]	50.81±15.16 (50.5) (n=16)	59±1.41 (59) (n=2)	[b]
Reduction of radial nerve CMAP amplitude in percentage (%) in the arm segment	68.14±29.77 (82.75) (n=21)	73.83±45.31 (100) (n=3)	[b]

CMAP: Compound muscle action potential, CNAP: Compound nerve action potential, ERNAS: Entrapment of the radial nerve at the spiral groove of humerus in the arm segment, DASH: The disabilities of the arm: shoulder and hand questionnaire, MRC: Medical research council, NCV: Nerve conduction velocity, SD: Standard deviation, TRNAS: Traumatic radial neuropathy in the arm segment, [a]: In some patients, NCV could not be calculated due to the absence of CNAP or distal CMAP, or because proximal CMAP could not be recorded as a result of motor conduction block, [b]: Because of the small sample size, statistical comparison was not performed. Values with p<0.05 are shown in bold

patients. Superficial radial nerve CNAP amplitudes and radial nerve CMAP amplitudes were found to be lower in the patients with TRNAS than in the patients with ERNAS.

The etiology of radial neuropathy includes conditions such as humeral fracture and systemic diseases (1,3,4). Another reason is ERNAS. This condition is also referred to as saturday night palsy (1,4,9). Similar to the characteristics of the patients in this current study, saturday night palsy can be seen after prolonged sleep as a result of alcohol or substance use, or after surgery (1,3,4). Demyelination

Table 2. Electrodiagnostic abnormalities of the patients

Abnormal electrodiagnostic features	Number of patients with abnormality of electrodiagnostic feature (%) n=30	Number of ERNAS patients with abnormality of electrodiagnostic feature (%) n=22	Number of TRNAS patients with abnormality of electrodiagnostic feature (%) n=8	Comparison between the ERNAS and TRNAS patients p-value
Superficial radial nerve CNAP amplitude	10 (33.3%)	4 (18.2%)	6 (75.0%)	**0.007**
Superficial radial nerve NCV	5[a,b] (20%)	5[a] (23.8%)	0[b] (0%)	0.549
Radial nerve CMAP amplitude	6 (20%)	1 (4.5%)	5 (62.5%)	**0.002**
Radial motor NCV across the arm segment	8[a,b] (44.4%)	8[a] (50%)	0[b] (0%)	0.526
Motor conduction block of radial nerve in the arm segment	17[a,b] (70.8%)	15[a] (71.4%)	2[b] (66.7%)	1.000
PSW and fibrillation potentials				
EIP	23 (76.6%)	17 (77.3%)	6 (75.0%)	1.000
EDC	28 (93.3%)	21 (95.5%)	7 (87.5%)	0.469
BRC	19 (63.3%)	15 (68.2%)	4 (50.0%)	0.390
Absence of MUAP				
EIP	7 (23.3%)	3 (13.6%)	4 (50.0%)	0.060
EDC	8 (26.6%)	4 (18.2%)	4 (50.0%)	0.158
BRC	5 (16.6%)	2 (9.1%)	3 (37.5%)	0.102

BRC: Brachioradialis muscle, CMAP: Compound muscle action potential, CNAP: Compound nerve action potential, EDC: Extensor digitorum communis, EIP: Extensor indicis proprius, ERNAS: Entrapment of the radial nerve at the spiral groove of humerus in the arm segment, MUAP: Motor unit action potential, NCV: Nerve conduction velocity, PSW: Positive sharp wave, TRNAS: Traumatic radial neuropathy in the arm segment, [a]: CNAP could not be obtained in one ERNAS patient/CMAP could not be obtained in one ERNAS patient/proximal CMAP could not be obtained in five ERNAS patients; [b]: CNAP could not be obtained in four TRNAS patients/CMAP could not be obtained in five TRNAS patients/proximal CMAP could not be obtained in one patient [accordingly, sensory NCV values reported for 25 patients (21 ERNAS, 4 TRNAS), and motor NCV values for 18 patients (16 ERNAS, 2 TRNAS)]. Values with p<0.05 are shown in bold

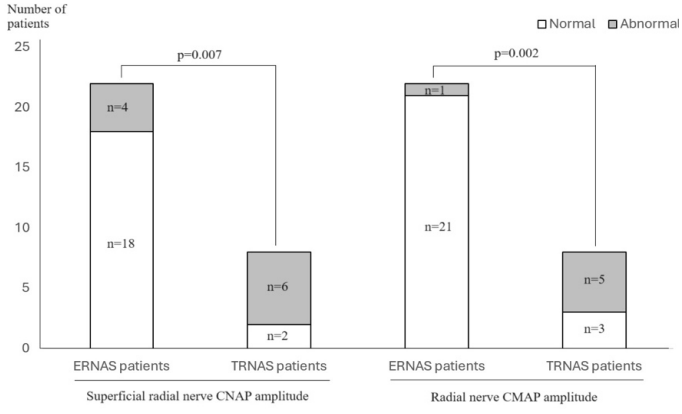


Figure 1. Amplitude abnormalities of superficial radial nerve CNAP and radial nerve CMAP among groups

CMAP: Compound muscle action potential, CNAP: Compound nerve action potential, ERNAS: Entrapment of the radial nerve at the spiral groove of humerus in the arm segment, TRNAS: Traumatic radial neuropathy in the arm segment

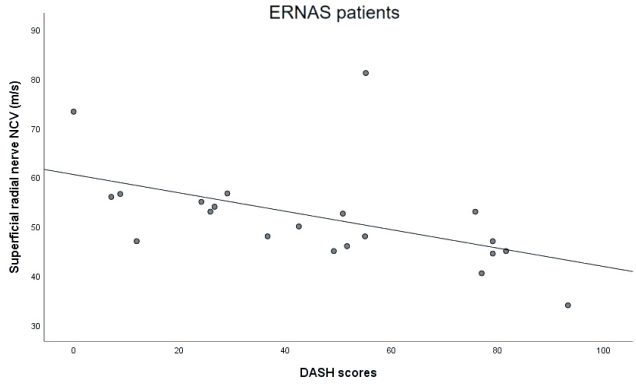


Figure 2. Negative correlation between superficial radial nerve NCV and DASH scores in ERNAS patients

DASH: The disabilities of the arm, shoulder and hand questionnaire, ERNAS: Entrapment of the radial nerve at the spiral groove of humerus in the arm segment, NCV: Nerve conduction velocity

Table 3. The correlation between DASH scores and electrodiagnostic findings in patients

Groups	SRN CNAP amplitude (μV)	SRN NCV (m/s)	RN CMAP amplitude (mV)	RN NCV across the arm segment	RN CMAP amplitude reduction in percentage in the arm segment (%)
DASH scores of ERNAS patients	p=0.960 r=0.011 (n=22)	**p=0.001,** **r=-0.677** (n=21)	p=0.713 r=0.083 (n=21)	p=0.312 r=-0.27 (n=21)	**p=0.044,** **r=-0.443** (n=21)
DASH scores of TRNAS patients	p=0.696 r=0.165 (n=8)	[a]	[a]	[a]	[a]
DASH scores of all patients	p=0.382, r=-0.165 (n=30)	**p=0.044,** **r=-0.406** (n=25)	p=0,136, r=-0.279 (n=24)	p=0.240, r=-0.292 (n=24)	p=0.137 r=0.312 (n=24)

CMAP: Compound muscle action potential, CNAP: Compound nerve action potential, DASH: The disabilities of the arm: shoulder and hand questionnaire, ERNAS: Entrapment of the radial nerve at the spiral groove of humerus in the arm segment, RN: Radial nerve, SRN: Superficial radial nerve, TRNAS: Traumatic radial neuropathy in the arm segment, [a]: Because of the small sample size, statistical comparison was not performed. Values with p<0.05 are shown in bold

contributes fundamentally to the pathophysiology of ERNAS patients (1,5,9). In radial neuropathy in the arm segment, patients present with a droopy hand (1,3,4). Sensory abnormalities in the distribution of the superficial radial nerve were noted in a substantial portion of patients, especially among those with ERNAS. This may be related to the demyelinating nature of ERNAS, where conduction impairment can occur without significant axonal loss. On the other hand, more pronounced sensory symptoms may develop in axonal lesions such as those seen in TRNAS. Nevertheless, since the difference between the groups did not reach statistical significance, these observations should be interpreted with caution. Previous studies have shown that demyelinating neuropathies generally have a more favorable prognosis than axonal injuries, particularly in terms of functional recovery (1,2,14).

Electrodiagnostic tests can provide important clues about the prognosis of RNIAS patients. The prognosis of nerve damage accompanied by axonal degeneration tends to be worse than the prognosis of nerve damage with demyelination alone (1,2,10,14). Recovery may be delayed or incomplete in axonal degeneration (1,2,5). Demyelination is predominant in ERNAS patients, and most ERNAS patients recover fully or almost completely in the first few months (1,5,9). Recovery rarely takes up to six months. In this current study, the radial nerve CMAP amplitudes of TRNAS patients were found to be lower than the CMAP amplitudes of ERNAS patients. This finding indicates that the type of radial nerve injury in ERNAS patients is compatible with segmental demyelination, and that the type of radial nerve injury in TRNAS patients is compatible with axonal degeneration. In light of this information, it can be said that the recovery in TRNAS patients may be later or worse than the recovery in ERNAS patients.

Radial nerve CMAP amplitude abnormality, radial motor nerve conduction block in the arm segment, radial motor NCV abnormality, and superficial radial nerve CNAP abnormality constitute the nerve conduction study findings in ERNAS and TRNAS patients (1,2,5). Motor conduction block and needle EMG abnormalities did not differ between ERNAS and TRNAS patients. As previously mentioned, there were electrodiagnostic differences between ERNAS and TRNAS patients in terms of abnormalities in radial nerve CMAP and superficial radial nerve CNAP amplitudes.

In this current study, a relationship was found between DASH scores and some electrodiagnostic findings. An inverse correlation was found between the severity of conduction block and DASH score in ERNAS patients. As the conduction block improves, the weakness may improve (15,16). Additionally, an inverse correlation was found between superficial radial nerve NCV and DASH scores of ERNAS patients and all RNIAS patients. Conduction block and superficial radial sensory NCV may have significant benefits in the follow-up of ERNAS patients. A reduced radial nerve CMAP amplitude in TRNAS patients may be associated with poor prognosis.

Study Limitations

There were some limitations in this study. Disease duration different significantly between ERNAS and TRNAS patients and should be considered an important potential confounder. In addition, the disease durations of patients within the same groups also differed. Electrodiagnostic test findings and clinical features may be influenced by disease duration (17). We think that this difference is caused by the fact that ERNAS patients are admitted in the acute period and TRNAS patients are initially presented with splints or injuries or are admitted after emergency surgical intervention. Despite this duration variability, we believe

that the observed differences in parameters such as CMAP and CNAP amplitudes reflect inherent differences in nerve involvement patterns, with demyelination being more common in ERNAS and axonal damage more prominent in TRNAS. Nevertheless, the potential impact of disease duration should be acknowledged when interpreting the results, and the findings should be interpreted with appropriate caution. The low number of TRNAS patients was also a limitation, but it should be kept in mind that the exclusion criteria in the current study were strict. However, studies that do not have strict exclusion criteria and include patients with diseases that can cause polyneuropathy, such as diabetes mellitus, or those with polyneuropathy may benefit from elucidating the pathophysiology and clinical features of radial neuropathy.

Conclusion

In summary, this study showed that radial nerve CMAP and superficial radial nerve CNAP amplitudes in TRNAS patients were lower than those in ERNAS patients. This may indicate that the prognosis in TRNAS patients is worse than in ERNAS patients. These findings should be interpreted with consideration of the differing disease durations between groups, as this may have influenced the degree of nerve injury observed on electrodiagnostic testing. It was found that there may be a relationship between superficial radial sensory NCV/radial motor nerve conduction block and clinical features in ERNAS patients. A similar relationship existed between radial nerve CMAP amplitude and clinical features in TRNAS patients.

Ethics

Ethics Committee Approval: Ethics committee approval was received from the Clinical Research Ethics Committee of University of Health Sciences Turkey, Adana City Training and Research Hospital (number: 118/2303, date: 15.12.2022).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: Ş.B., H.F., H.C.A., Concept: Ş.B., H.F., Design: Ş.B., H.F., Data Collection or Processing: Ş.B., H.F., H.C.A., Z.A., Analysis or Interpretation: Ş.B., H.F., H.C.A., Z.A., Literature Search: Ş.B., H.F., H.C.A., Z.A., Writing: Ş.B., H.F., H.C.A., Z.A.

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

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

References

1. Doughty CT, Bowley MP. Entrapment neuropathies of the upper extremity. Med Clin North Am. 2019;103(2):357-370.
2. Wang LH, Weiss MD. Anatomical, clinical, and electrodiagnostic features of radial neuropathies. Phys Med Rehabil Clin N Am. 2013;24(1):33-47.
3. Węgiel A, Karauda P, Zielinska N, Tubbs RS, Olewnik L. Radial nerve compression: anatomical perspective and clinical consequences. Neurosurg Rev. 2023;46(1):53.
4. Vij N, Kiernan H, Miller-Gutierrez S, Agusala V, Kaye AD, Imani F, et al. Etiology diagnosis and management of radial nerve entrapment. Anesth Pain Med. 2021;11(1):e112823.
5. Kim KH, Park KD, Chung PW, Moon HS, Kim YB, Yoon WT, et al. The usefulness of proximal radial motor conduction in acute compressive radial neuropathy. J Clin Neurol. 2015;11(2):178-182.
6. Kleyweg R, van der Meche F, Schimtz P. Interobserver agreement in the assessment of muscle strength and functional abilities in Guillain-Barré syndrome. Muscle Nerve. 1991;14(11):1103-1109.
7. Hudak PL, Amadio PC, Bombardier C. Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). Am J Ind Med. 1996;29(6):602-608. doi: 10.1002/(SICI)1097-0274(199606)29:6<602::AID-AJIM4>3.0.CO;2-L. Erratum in: Am J Ind Med 1996;30(3):372.
8. Düger T, Yakut E, Öksüz C, Yörükan S, Bilgutay BS, Ayhan Ç, et al. Reliability and validity of the Turkish version of the disabilities of the arm, shoulder and hand (DASH) questionnaire. Turk J Physiother Rehabil 2006;17(3):99-107.
9. Arnold WD, Krishna VR, Freimer M, Kisset JT, Elsheikh B. Prognosis of acute compressive radial neuropathy. Muscle Nerve. 2012;45(6):893-894.
10. Carlson N, Logigian EL. Radial neuropathy. Neurol Clin. 1999;17(3):499-523.
11. Oh SJ. Clinical electromyography nerve conduction studies. 3rd ed. Philadelphia: Lippincott Williams & Wilkins;2003:231-234.
12. Chen S, Andary M, Buschbacher R, Del Toro D, Smith B, So Y, et al. Electrodiagnostic reference values for upper and lower limb nerve conduction studies in adult populations. Muscle Nerve. 2016;54(3):371-377.
13. Fidancı H, Öztürk İ, Köylüoğlu AC, Yildiz M, Buturak Ş, Arlier Z. The needle electromyography findings in the neurophysiological classification of ulnar neuropathy at the elbow. Turk J Med Sci. 2020;50(4):804-810.
14. Sunderland S. A classification of peripheral nerve injuries producing loss of function. Brain. 1951;74(4):491-516.
15. Capasso M, Caporale CM, Pomilio P, Gandolfi P, Lugaresi A, Uncini A. Acute motor conduction block neuropathy another Guillain-Barré syndrome variant. Neurology. 2003;61(5):617-22.
16. Kokubun N, Nishibayashi M, Uncini A, Odaka M, Hirata K, Yuki N. Conduction block in acute motor axonal neuropathy. Brain. 2010;133(10):2897-2898.
17. Bergquist ER, Hammert WC. Timing and appropriate use of electrodiagnostic studies. Hand Clin. 2013;29(3):363-370.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):213-218
DOI: 10.4274/BMB.galenos.2026.41275

Investigation of the Effect of Multi-screen Addiction on Adult Individuals Diagnosed with Dry Eye

Göz Kuruluğu Tanısı Konulan Yetişkin Bireylerde Çoklu Ekran Bağımlılığının Etkisinin Araştırılması

Güney Yıldız[1], Sibel Tunç Karaman[1], Umut Şahin[2], Okcan Basat[1]

[1]University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, Department of Family Medicine, İstanbul, Turkey
[2]University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, Department of Ophthalmology, İstanbul, Turkey

Abstract

Objective: Dry eye is an ocular surface disease characterized by the loss of the tear film layer. Alongside the increasing use of digital screen devices, the prevalence of dry eye has also risen. This study aims to investigate the effect of multi-screen addiction on dry eye in adults.

Method: This cross-sectional study was conducted between 01.05.2023 and 01.11.2023 with adults who presented to the Family Medicine and Ophthalmology Outpatient Clinics of a University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital with symptoms of dry eye and met the inclusion criteria. All participants were evaluated for dry eye using the Schirmer test. According to the tear film and ocular surface society dry eye workshop II criteria, 210 individuals were classified as having dry eye, while 210 individuals were classified as not having dry eye. Data were collected using a demographic information form and the multi-screen addiction scale (MSAS). A p-value of <0.05 was considered statistically significant.

Results: A total of 420 individuals with a mean age of 37.35±13.58 years (range: 18-79) were included in the study. The mean total score on the MSAS for all participants was 36.16±11.94 (range: 15-67). According to the MSAS, 41.4% (n=174) of all participants exhibited multi-screen addiction, with rates of 50.0% (n=105) in the dry eye group and 32.9% (n=69) in the non-dry eye group. The prevalence of multi-screen addiction was statistically significantly higher among individuals with dry eye compared to those without dry eye (p<0.001).

Öz

Amaç: Göz kuruluğu, gözyaşı film tabakasının kaybı ile karakterize bir oküler yüzey hastalığıdır. Dijital ekranlı cihazların kullanımının artmasına paralel olarak göz kuruluğu görülme sıklığı da artmaktadır. Bu çalışmada; yetişkinlerde çoklu ekran bağımlılığının göz kuruluğuna etkisinin araştırılması amaçlanmıştır.

Yöntem: Bu kesitsel çalışma; 01.05.2023-01.11.2023 tarihleri arasında Sağlık Bilimleri Üniversitesi, Gaziosmanpaşa Eğitim ve Araştırma Hastanesi, Aile Hekimliği ve Göz Hastalıkları Poliklinikleri'ne göz kuruluğu semptomlarıyla başvuran ve çalışmaya dahil edilme kriterlerini karşılayan yetişkinler ile gerçekleştirildi. Tüm katılımcılar Schirmer testi ile göz kuruluğu açısından değerlendirildi. Gözyaşı filmi ve oküler yüzey topluluğu kuru göz atölyesi kriterlerine göre 210 katılımcı göz kuruluğu olan olarak sınıflandırılırken, 210 katılımcı göz kuruluğu olmayan olarak sınıflandırıldı. Verileri elde etmede tanıtıcı bilgi formu ve çoklu ekran bağımlılığı ölçeği (ÇEBÖ) kullanıldı. P<0,05 istatistiksel olarak anlamlı kabul edildi.

Bulgular: Çalışmaya yaş ortalaması 37,35±13,58 (min: 18-maks: 79) yıl olan toplam 420 kişi dahil edildi. Tüm katılımcıların ÇEBÖ'den aldıkları toplam puan ortalama 36,16±11,94 (min: 15-maks: 67) idi. ÇEBÖ'ye göre tüm katılımcıların %41,4'ünde (n=174) çoklu ekran bağımlılığı saptandı; bu oran göz kuruluğu olan grupta %50,0 (n=105), göz kuruluğu olmayan grupta ise %32,9 (n=69) idi. Çoklu ekran bağımlılığı olma oranı göz kuruluğu olanlarda göz kuruluğu olmayanlara göre istatistiksel olarak anlamlı şekilde daha yüksekti (p<0,001).

Address for Correspondence: Sibel Tunç Karaman, University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital, Department of Family Medicine, İstanbul, Turkey
E-mail: drsibeltunc@hotmail.com **ORCID:** orcid.org/0000-0003-1833-8758
Received: 28.05.2025 **Accepted:** 03.06.2026 **Epub:** 08.06.2026 **Publication Date:** 24.06.2026

Cite this article as: Yıldız G, Tunç Karaman S, Şahin U, Basat O. Investigation of the effect of multi-screen addiction on adult individuals diagnosed with dry eye. Bagcilar Med Bull. 2026;11(2):213-218

Abstract

Conclusion: In this study, multi-screen addiction was identified in 41.4% of participants. The prevalence of multi-screen addiction was statistically significantly higher among individuals with dry eye (50.0%) compared to those without dry eye (32.9%). These findings suggest that behavioral factors related to screen use may be associated with ocular surface health. Therefore, monitoring screen time and promoting preventive eye health measures, particularly among individuals with intensive digital screen use, are considered important.

Keywords: Dry eye, family medicine, Schirmer test, screen addiction

Öz

Sonuç: Bu çalışmada; katılımcıların %41,4'ünde çoklu ekran bağımlılığı saptandı. Çoklu ekran bağımlılığı prevalansı, göz kuruluğu olan bireylerde (%50,0) göz kuruluğu olmayanlara (%32,9) kıyasla istatistiksel olarak anlamlı derecede daha yüksekti. Bu bulgular, ekran kullanımına ilişkin davranışsal faktörlerin oküler yüzey sağlığı ile ilişkili olabileceğini düşündürmektedir. Dolayısıyla, özellikle yoğun dijital ekran kullanımı olan bireylerde ekran süresinin izlenmesi ve koruyucu göz sağlığı önlemlerinin teşvik edilmesi önem taşımaktadır.

Anahtar kelimeler: Aile hekimliği, ekran bağımlılığı, göz kuruluğu, Schirmer testi

Introduction

According to the tear film and ocular surface society dry eye workshop II (TFOS DEWS II) reports, dry eye is defined as a multifactorial disease characterized by a loss of homeostasis of the tear film, accompanied by ocular symptoms. In the etiopathogenesis of dry eye, tear film instability and hyperosmolarity, ocular surface inflammation and damage, as well as neurosensory abnormalities, play significant roles (1-3). Dry eye may cause symptoms such as burning, itching, redness, and ocular fatigue. Blurred vision and photophobia are also common, particularly during visually demanding tasks. If the condition persists over an extended period, it can interfere with daily activities and may lead to permanent ocular damage (4). The global prevalence of dry eye is estimated to range between 5% and 50%, varying based on the demographic and clinical characteristics of the studied populations (5). Epidemiological studies have reported that the prevalence of dry eye increases with age and is more common among women (6,7). Moreover, office workers, students, and individuals with prolonged exposure to digital screens are considered high-risk groups (8,9).

In recent years, the widespread use of digital screen devices—such as computers, televisions, tablets, and smartphones—has made dry eye an increasingly important public health concern. Prolonged screen time can disrupt the tear film and reduce the blink rate, thereby contributing to the development of dry eye. Increased exposure to screens accelerates tear evaporation, particularly promoting the onset of evaporative dry eye (10).

Excessive use of digital screen devices has also been associated with various health issues, including screen addiction. Screen addiction comprises several subtypes, such as internet addiction, digital gaming addiction, media addiction, and technological device addiction (11,12). Although multi-screen use offers certain advantages in daily life—such as rapid access to information and multifaceted interaction—its excessive use may lead to addiction and a range of adverse health outcomes, including sleep disturbances, physical inactivity, overeating, weight gain, and obesity (13-18).

Current research on health problems associated with screen addiction primarily focuses on systemic effects. However, the relationship between multi-screen addiction and dry eye has not been adequately explored in the literature. Therefore, this study aims to evaluate the impact of multi-screen addiction on dry eye.

Materials and Methods

Study Design and Participants

This study was designed as a cross-sectional study. It was conducted between May 1, 2023, and November 1, 2023, among individuals presenting with symptoms suggestive of dry eye at the Family Medicine and Ophthalmology Outpatient Clinics of a University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital.

Sample size calculation was performed using G*Power software. With a statistical power of 95%, a significance level of 5%, and an effect size of 0.26, the minimum required sample size was calculated as 194 participants in each group (dry eye and non-dry eye). A total of 420 participants were ultimately enrolled in the study.

Eligibility Criteria

Inclusion Criteria

- Age ≥18 years,
- Presentation with at least one symptom suggestive of dry eye (burning, itching, redness, ocular fatigue, blurred vision, photophobia),
- Ability to provide informed consent,
- Completion of Schirmer test and study questionnaires.

Exclusion Criteria

- Presence of an acute illness at the time of evaluation,
- Systemic diseases strongly associated with dry eye (e.g., Sjögren's syndrome),
- Use of medications affecting sleep or alertness (e.g., benzodiazepines, antipsychotics, H1 antihistamines),
- Hearing, speech, cognitive, or cooperation limitations impairing communication,
- Pregnancy.

Data Collection

Schirmer Test Procedure

During the initial assessment, all participants were evaluated for dry eye by an ophthalmologist using the Schirmer test. The Schirmer test was performed as Schirmer I without anesthesia, by placing standardized Schirmer strips on the outer one-third of the lower conjunctival fornix of both eyes for 5 minutes to measure basal and reflex tear production. Based on the TFOS DEWS II criteria, participants with tear production <10 mm/5 min were classified as having dry eye and included in the dry eye group, whereas those with tear production >10 mm/5 min were assigned to the non-dry eye group (3).

Descriptive Information Form

The form developed by us included questions regarding the participants' socio-demographic characteristics (age, gender, education level), presence of chronic diseases and use of eyeglasses.

Multiple Screen Addiction Scale (MSAS)

The MSAS, developed by Sarıtepeci (19) in 2021, is a 15-item, 5-point Likert-type scale used to assess multi-screen addiction. The scale comprises three subdimensions: Excessive screen time (4 items), compulsive behavior [(CB) 8 items], and loss of control [(LC) 3 items]. The MSAS defines addiction according to the monothetic criterion as scoring 3 or above on all items, and according to the polythetic criterion as scoring 3 or above on at least 8 items. The Cronbach's alpha coefficients for the total scale and its subdimensions range between 0.71 and 0.92, indicating good internal consistency (19).

In the present study, the polythetic criterion (≥3 on at least 8 items) was applied to classify participants as having multi-screen addiction. Participants meeting this threshold were categorized as having addiction (yes/no), and this categorical variable was used in group comparisons. Total and subscale scores were additionally analyzed as continuous variables.

Ethical Considerations

The study was conducted with the approval of the Clinical Research Ethics Committee of University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital (date: 12.04.2023, approval number: 53). All procedures were carried out in accordance with the Declaration of Helsinki. Verbal and written informed consent were obtained from all participants.

Statistical Analyses

Statistical analyses of the data obtained in the study were performed using IBM SPSS Statistics for Windows, Version 25.0 (Statistical Package for the Social Sciences, IBM Corp., Armonk, NY, USA). Descriptive statistics were presented as frequency (n) and percentage (%) for categorical variables, and mean ± standard deviation for continuous variables. Normality assumptions were assessed using the Kolmogorov-Smirnov test, with p-values >0.05 indicating normal distribution. Independent samples t-test was used to compare various clinical parameters between groups. Pearson's chi-square test was employed for comparisons of categorical variables. A p-value of less than 0.05 was considered statistically significant.

Results

This study was conducted with a total of 420 participants, including 210 individuals with dry eye and 210 without dry eye. The mean age of all participants was 37.35±13.58 years (range: 18-79). The mean age of the dry eye group was 36.51±13.45 years (range: 18-75), and the non-dry eye group had a mean age of 38.17±13.69 years (range: 18-79). The frequency distributions of the socio-demographic and medical characteristics of the two groups are presented in Table 1.

Descriptive statistics of the MSAS scores and the distribution of multi-screen addiction according to dry eye status are presented in Table 2. The mean total MSAS score of all participants was 36.16±11.94 (range: 15-67). Participants with dry eye demonstrated higher total MSAS scores (37.83±12.47) compared to those without dry eye (34.49±11.18). Similarly, subscale scores for excessive screen time, CB, and LC were higher in the dry eye group.

According to the polythetic classification criterion, 41.4% (n=174) of the total sample met the criteria for multi-screen addiction. The prevalence of multi-screen addiction was

Table 1. Socio-demographic and medical characteristics of the dry eye and non-dry eye groups

	Total (n=420)	Dry eye (-) (n=210)	Dry eye (+) (n=210)
Age (mean ± SD)	37.35±13.58	38.17±13.69	36.51±13.45
	n (%)	n (%)	n (%)
Gender			
Male	146 (34.8)	74 (35.2)	72 (34.3)
Female	274 (65.2)	136 (64.8)	138 (65.7)
Education level			
Primary and middle school	66 (15.7)	32 (15.2)	34 (16.2)
High school	115 (27.4)	66 (31.4)	49 (23.3)
Associate degree	48 (11.4)	13 (6.2)	35 (16.7)
University	191 (45.5)	99 (47.1)	92 (43.8)
Chronic disease			
No	312 (74.3)	150 (71.4)	162 (77.1)
Yes	108 (25.7)	60 (28.6)	48 (22.9)
Use of eyeglasses			
No	221 (52.6)	98 (46.7)	123 (58.6)
Yes	199 (47.4)	112 (53.3)	87 (41.4)

Data are presented as n (%) and mean ± SD values. SD: Standard deviation

Table 2. Descriptive statistics of MSAS scores and distribution of multi-screen addiction according to dry eye status

	Min-max	Mean ± SD
Dry eye (+)		
Total screen addiction score	15.00-67.00	37.83±12.47
Excessive screen time	4.00-20.00	11.08±3.69
Compulsive behavior	8.00-40.00	20.58±7.42
Loss of control	2.00-14.00	6.17±2.99
Dry eye (-)		
Total screen addiction score	18.00-66.00	34.49±11.18
Excessive screen time	4.00-20.00	9.72±4.17
Compulsive behavior	9.00-37.00	18.90±6.68
Loss of control	3.00-14.00	5.88±2.19
Screen addiction according to MSAS (polythetic criterion)		
	No	**Yes**
	n (%)	**n (%)**
Dry eye (-)	141 (67.1)	69 (32.9)
Dry eye (+)	105 (50.0)	105 (50.0)
Total	246 (58.6)	174 (41.4)

Data are presented as n (%), min-max and mean ± SD values. SD: Standard deviation, MSAS: Multi-screen addiction scale

50.0% among individuals with dry eye and 32.9% among those without dry eye.

As shown in Table 3, there was a statistically significant difference between the dry eye and non-dry eye groups regarding multi-screen addiction status (p<0.001). The proportion of participants classified as having multi-screen addiction was significantly higher in the dry eye group compared to the non-dry eye group.

In addition, dry eye status was significantly associated with eyeglass use (p=0.015) and education level (p=0.005). No statistically significant differences were observed between groups with respect to age, gender, marital status, or presence of chronic disease (p>0.05).

Discussion

This study examined the relationship between multi-screen addiction and dry eye by comparing individuals with and without dry eye. Multi-screen addiction was identified in 41.4% of participants overall. The prevalence of multi-screen addiction was significantly higher in the dry eye group (50.0%) compared to the non-dry eye group (32.9%).

Table 3. Comparison of socio-demographic and clinical characteristics according to dry eye status

	Dry eye (-) (n=210)	Dry eye (+) (n=210)	p
Age (mean ± SD)	38.18±13.70	36.52±13.45	0.212[a]
Gender	**n (%)**	**n (%)**	
Male	74 (35.2)	72 (34.3)	0.838[b]
Female	136 (64.8)	138 (65.7)	
Marital status			
Married	112 (53.3)	123 (58.6)	0.280[b]
Single	98 (46.7)	87 (41.4)	
Use of eyeglasses			
No	98 (46.7)	123 (58.6)	0.015[b]
Yes	112 (53.3)	87 (41.4)	
Education level			
Primary and middle school	32 (15.2)	34 (16.2)	0.005[b]
High school	66 (31.4)	49 (23.3)	
Associate degree	13 (6.2)	35 (16.7)	
University	99 (47.1)	92 (43.8)	
Chronic disease			
No	150 (71.4)	162 (77.1)	0.180[b]
Yes	60 (28.6)	48 (22.9)	
Screen addiction according to MSAS			
Yes	69 (32.9)	105 (50.0)	<0.001[b]

SD: Standard deviation, MSAS: Multi-screen addiction scale, [a]: Independent t-test, [b]: Pearson chi-square test, p<0.05 is statistically significant

Dry eye is a multifactorial ocular surface disorder characterized by tear film instability. With the widespread and prolonged use of digital devices in occupational, educational, and recreational settings, the frequency of dry eye-related symptoms has increased markedly (20). Several studies have demonstrated an association between prolonged screen use and dry eye (21,22). Portello et al. (23) showed that extended exposure to computers and other screen-based devices reduces blink rate, increasing tear evaporation and contributing to dry eye development. Evidence also indicates a strong correlation between longer screen use and the prevalence and severity of symptoms (24). Tallens-Estarelles et al. (25) similarly reported increased dry eye signs among screen users, highlighting reduced tear volume, impaired stability, and altered tear film composition as key mechanisms. Muntz et al. (26) found that individuals with excessive weekly screen time frequently exhibited dry eye symptoms, with 90% being symptomatic according to the SANDE questionnaire, and demonstrated a positive correlation between prolonged screen exposure, compensatory blink changes, and reduced tear film stability.

In our study, dry eye was assessed using the Schirmer I test without anesthesia, and no subtype classification was performed. Multi-screen addiction was significantly more frequent among those with dry eye (50.0%). These findings reinforce the association between screen-related behaviors and ocular surface complaints, emphasizing the role of behavioral factors in the increasing clinical burden of dry eye. Understanding the etiological contribution of multi-screen addiction is therefore essential, as reducing excessive screen exposure may help mitigate both the prevalence and severity of dry eye symptoms.

Chronic systemic diseases such as diabetes, autoimmune disorders, and thyroid dysfunction are recognized contributors to ocular surface instability. However, in our study, the prevalence of chronic disease did not differ significantly between groups. TFOS DEWS II reports note that systemic comorbidities can disrupt tear film homeostasis (1), yet the distribution of chronic conditions in our sample largely involved disorders not strongly associated with tear film abnormalities. This may explain the absence of a significant relationship, suggesting that behavioral factors—particularly prolonged screen exposure—may have played a more prominent role in dry eye prevalence in this population.

In addition to these findings, dry eye was more common among participants who did not use eyeglasses. Although eyeglasses are not a treatment modality, TFOS DEWS II reports indicate that reduced ocular surface protection and environmental exposure can increase tear film instability and evaporation (1). By providing partial shielding, eyeglasses may help reduce evaporative stress, and individuals who do not use them may experience greater ocular surface exposure during prolonged screen viewing (25). Although our study assessed general eyeglass use rather than sunglasses specifically, prior evidence suggests that various forms of optical or protective eyewear, including sunglasses, can reduce airflow exposure and tear evaporation, which may partially explain our findings.

Another noteworthy finding was the higher prevalence of dry eye among participants with higher educational attainment. This group is more likely to engage in prolonged digital device use, a recognized risk factor for dry eye (21). Prior studies show that intensive screen exposure reduces blink rate and destabilizes the tear film (10,23), which may explain the increased dry eye prevalence observed in this subgroup.

Study Limitations

This study has several limitations. Firstly, dry eye was assessed using the Schirmer I test without anesthesia, without further subtype classification, which limited the analysis of subtype-specific relationships with multi-screen addiction. Secondly, the cross-sectional design prevents establishing a causal relationship between multi-screen addiction and dry eye. Future studies should include detailed classification of dry eye subtypes and further explore their associations with multi-screen addiction. Longitudinal studies are also needed to clarify the long-term effects of screen use on ocular health.

Conclusion

The present study highlights the potential impact of the increasing multi-screen use on eye health. Multi-screen addiction was common among participants (41.4%) and significantly higher in those with dry eye (50.0%). The findings suggest that lifestyle and individual factors contribute to dry eye development, with increased screen time causing adverse changes to the ocular surface. Therefore, it is essential for individuals to monitor and manage their screen time and follow preventive measures to protect their eyes.

Ethics

Ethics Committee Approval: The study was conducted with the approval of the Clinical Research Ethics Committee of University of Health Sciences Turkey, Gaziosmanpaşa Training and Research Hospital (date: 12.04.2023, approval number: 53). All procedures were carried out in accordance with the Declaration of Helsinki.

Informed Consent: Verbal and written informed consent were obtained from all participants.

Footnotes

Authorship Contributions

Surgical and Medical Practices: G.Y., U.Ş., Concept: G.Y., U.Ş., O.B., Design: G.Y., S.T.K., U.Ş., O.B., Data Collection or Processing: G.Y., S.T.K., U.Ş., Analysis or Interpretation: G.Y., S.T.K., U.Ş., O.B., Literature Search: G.Y., S.T.K., Writing: G.Y., S.T.K., O.B.

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

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

References

1. Craig JP, Nichols KK, Akpek EK, Caffery B, Dua HS, Joo CK, et al. TFOS DEWS II definition and classification report. Ocul Surf. 2017;15(3):276-283.
2. Jones L, Downie LE, Korb D, Benitez-Del-Castillo JM, Dana R, Deng SX, et al. TFOS DEWS II management and therapy report. Ocul Surf. 2017;15(3):575-628.
3. The definition and classification of dry eye disease: report of the definition and classification subcommittee of the International Dry Eye WorkShop (2007). Ocul Surf. 2007;5(2):75-92
4. Sheppard J, Shen Lee B, Periman LM. Dry eye disease: identification and therapeutic strategies for primary care clinicians and clinical specialists. Ann Med. 2023;55(1):241-252.
5. Stapleton F, Alves M, Bunya VY, Jalbert I, Lekhanont K, Malet F, et al. TFOS DEWS II epidemiology report. Ocul Surf. 2017;15(3):334-365.
6. Farrand KF, Fridman M, Stillman IÖ, Schaumberg DA. Prevalence of diagnosed dry eye disease in the United States among adults aged 18 years and older. Am J Ophthalmol. 2017;182:90-98.
7. Alkabbani S, Jeyaseelan L, Rao AP, Thakur SP, Warhekar PT. The prevalence, severity, and risk factors for dry eye disease in Dubai - a cross sectional study. BMC Ophthalmol. 2021;21(1):219.
8. Van Tilborg MM, Murphy PJ, Evans KS. Impact of dry eye symptoms and daily activities in a modern office. Optom Vis Sci. 2017;94(6):688-693.
9. Al-Mohtaseb Z, Schachter S, Shen Lee B, Garlich J, Trattler W. The relationship between dry eye disease and digital screen use. Clin Ophthalmol. 2021;15:3811-3820.
10. Akkaya S, Atakan T, Acikalin B, Aksoy S, Ozkurt Y. Effects of long-term computer use on eye dryness. North Clin Istanb. 2018;5(4):319.
11. Karakose T, Tülübaş T, Papadakis S. Revealing the intellectual structure and evolution of digital addiction research: an integrated bibliometric and science mapping approach. Int J Environ Res Public Health. 2022;19(22):14883.
12. Lopez-Fernandez O. Generalised versus specific internet use-related addiction problems: a mixed methods study on internet, gaming, and social networking behaviours. Int J Environ Res Public Health. 2018;15(12):2913.
13. Sa'ari AS, Hamid MRA, 'Ain Azizan N, Ismail NH. Examining the evidence between screen time and night eating behaviour with dietary intake related to metabolic syndrome: a narrative review. Physiol Behav. 2024;280:114562.
14. Aragay N, Vallès V, Ramos-Grille I, Garrido G, Grimalt EG, Miranda Ruiz E, et al. Differences in screen addiction in the past 15 years. Int J Environ Res Public Health. 2023;21(1):1.
15. Rao SR, Mhatre PP, Ambike AA, Adarkar SA, Gupta V, Paranjape RM, et al. Health impact of screen addiction: a cross-sectional comparative study. Natl Med J India. 2024;37(1):18-21.
16. Das P, Saraswathy KN, Chaudhary V. Prevalence of smartphone addiction and its relationship with obesity among young adults: a cross-sectional study from Delhi, India. Indian J Community Med. 2024;49(3):544-548.
17. Wong HY, Mo HY, Potenza MN, Chan MNM, Lau WM, Chui TK, et al. Relationships between severity of internet gaming disorder, severity of problematic social media use, sleep quality and psychological distress. Int J Environ Res Public Health. 2020;17(6):1879.
18. Burén J, Nutley SB, Thorell LB. Screen time and addictive use of gaming and social media in relation to health outcomes. Front Psychol. 2023;14:1258784.
19. Sarıtepeci M. Multiple screen addiction scale: validity and reliability study. Instructional Technology and Lifelong Learning. 2021;2(1):1-17.
20. Condori-Meza IB, Dávila-Cabanillas LA, Challapa-Mamani MR, Pinedo-Soria A, Torres RR, Yalle J, et al. Problematic internet use associated with symptomatic dry eye disease in medical students from Peru. Clin Ophthalmol. 2021;15:4357-4365.
21. Hanyuda A, Sawada N, Uchino M, Kawashima M, Yuki K, Tsubota K, et al. Physical inactivity, prolonged sedentary behaviors, and use of visual display terminals as potential risk factors for dry eye disease: JPHC-NEXT study. Ocul Surf. 2020;18(1):56-63.
22. Wang MTM, Muntz A, Mamidi B, Wolffsohn JS, Craig JP. Modifiable lifestyle risk factors for dry eye disease. Cont Lens Anterior Eye. 2021;44(6):101409.
23. Portello JK, Rosenfield M, Chu CA. Blink rate, incomplete blinks and computer vision syndrome. Optom Vis Sci. 2013;90(5):482-487.
24. Jakhar F, Rodrigues GR, Mendonca TM, Nayak RR, Kamath G, Kamath SJ, et al. Dry eye symptoms and digital eyestrain - emerging epidemics among university students due to online curriculum amid the COVID-19 pandemic. A cross-sectional study. Indian J Ophthalmol. 2023;71(4):1472-1477.
25. Talens-Estarelles C, García-Marqués JV, Cervino A, García-Lázaro S. Use of digital displays and ocular surface alterations: a review. Ocul Surf. 2021;19:252-265.
26. Muntz A, Turnbull PR, Kim AD, Gokul A, Wong D, Tsay TS, et al. Extended screen time and dry eye in youth. Cont Lens Anterior Eye. 2022;45(5):101541.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):219-226
DOI: 10.4274/BMB.galenos.2025.84856

The Relationship Between Women's Perspectives on Birth Methods and Their Socio-demographic Characteristics

Kadınların Doğum Şekillerine Olan Bakış Açıları ve Sosyo-demografik Özellikleri ile İlişkisi

Sibel Baktır Altuntaş[1], Suat Karataş[2]

[1]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Family Medicine, İstanbul, Turkey
[2]Private Nişantaşı Hospital, Clinic of Obstetrics and Gynecology, İstanbul, Turkey

Abstract

Objective: The World Health Organization recommends vaginal delivery while aiming for a cesarean delivery rate of 15% or lower. Factors influencing birth method preference include previous birth experiences and experiences during and after childbirth. Our study aimed to examine the past birth experiences of women who had previously given birth and to identify factors influencing birth method preferences.

Method: A descriptive cross-sectional study, designed as a questionnaire survey, was conducted with 434 participants at the Bahçelievler Memorial Hospital Obstetrics and Gynecology Polyclinic. Socio-demographic characteristics, past birth experiences, and questions related to maternal and infant health were administered face-to-face to voluntary participants. Data analysis was performed using SPSS 25 software. A p-value of <0.05 was considered statistically significant.

Results: As of the date of the study, the average number of vaginal deliveries was 2.48±1.13 (min: 1.0, max: 5.0), while the average number of cesarean deliveries was 1.87±0.79 (min: 1.0, max: 4.0). Reasons for cesarean delivery included stalled labor in 38.0% (n=113) and a history of previous cesarean delivery in 29.3% (n=87). When asked, "If you were to become pregnant again, which delivery method would you prefer?" 73.1% (n=321) answered vaginal delivery, while 26.9% (n=118) answered cesarean delivery. Having had a vaginal delivery in the first birth [odds ratio (OR) =6.54; 95% confidence interval (CI): 3.94-10.85; p<0.001] and the absence of a history of chronic

Öz

Amaç: Dünya Sağlık Örgütü, vajinal yolla yapılan doğumu önerirken sezaryen doğum oranının %15 ve altında olmasını istemektedir. Doğum şekli tercihinde birden fazla faktör etkendir. Çalışmamız ile polikliniğe başvuran daha önce doğum yapmış kadınların geçmiş doğum deneyimlerinin irdelenmesi ve doğum şekli tercihlerinde etkili olan faktörleri öğrenmeyi amaçladık.

Yöntem: Tanımlayıcı kesitsel, bir anket çalışması olarak planlanan çalışma Bahçelievler Memorial Hastanesi Kadın Doğum Polikliniği'nde 434 katılımcıyla yapıldı. Sosyo-demografik özellikleri, geçmiş doğum deneyimleri ve anne bebek sağlığı ile ilgili hazırlanan anket soruları çalışmaya katılmaya gönüllü olanlara yüz yüze uygulandı. Verilerin analizinde SPSS 25 programı kullanıldı. P<0,05 istatistiksel olarak anlamlı kabul edildi.

Bulgular: Çalışmanın yapıldığı tarihe kadar vajinal yolla doğum yapma sayısı ortalaması 2,48±1,13 (min: 1,0, maks: 5,0) iken sezaryen yolla doğum yapma sayısı ortalaması 1,87±0,79 (min: 1,0, maks: 4,0) idi. Sezaryen doğum yapma nedeni olarak %38,0 (n=113) ilerlemeyen doğum eylemi ve ardından %29,3 (n=87) daha önce sezaryen yolla doğum yapmış olmak geliyordu. "Bir gebelik yaşayacak olsanız, hangi doğum yöntemini tercih ederdiniz?" sorusuna (n=321) %73,1 vajinal doğum derken %26,9 (n=118) sezaryen doğum şekli diye cevapladı. İlk doğumun vajinal yol ile gerçekleşmiş olması [olasılık oranı (OO) =6,54; %95 güven aralığı (GA): 3,94-10,85; p<0,001] ve kronik hastalık öyküsünün bulunmaması (OO =3,68; %95 GA: 1,86-7,28; p<0,001) vajinal doğum tercihini etkileyen faktörlerdir.

Address for Correspondence: Sibel Baktır Altuntaş, University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Family Medicine, İstanbul, Turkey
E-mail: drsibell@gmail.com **ORCID:** orcid.org/0000-0003-1392-9014
Received: 03.06.2025 **Accepted:** 15.11.2025 **Epub:** 05.12.2025 **Publication Date:** 24.06.2026

Cite this article as: Baktır Altuntaş S, Karataş S. The relationship between women's perspectives on birth methods and their socio-demographic characteristics. Bagcilar Med Bull. 2026;11(2):219-226

Abstract

illness (OR =3.68; 95% CI: 1.86-7.28; p<0.001) were significant factors influencing the preference for vaginal delivery.

Conclusion: The previous mode of delivery influences the decision-making process regarding birth method. Women with a history of cesarean delivery prefer vaginal delivery. Experiences during and after previous deliveries are influential in deciding the method of childbirth.

Keywords: Cesarian section, childbirth choices, obstetric delivery, pregnancy

Öz

Sonuç: Doğum şekline karar verme üzerinde daha önce yapılan doğumun şekli etkilidir. Daha önce vajinal doğum hikayesi olan kadınlar vajinal yolla doğum yapmayı tercih etmektedir. Önceki doğum sırasında ve sonrasında yaşananlar doğum şekline karar vermede etkilidir.

Anahtar kelimeler: Doğum, doğum tercihi, hamilelik, sezaryen doğum

Introduction

Childbirth is a multifaceted event that encompasses physical, cultural, and medical dimensions. Vaginal delivery, as one of the physiological modes of birth, generally results in favorable outcomes when the mother is both physically and emotionally prepared. The alternative mode of delivery, cesarean section, is typically preferred in cases of cephalopelvic disproportion, fetal malpresentation, fetal distress, dystocia, macrosomia, or a history of cesarean delivery. Maternal morbidity and mortality rates associated with cesarean sections are reported to be 2 to 7 times higher than those associated with vaginal deliveries (1,2).

In our country, cesarean section—although frequently performed—is often perceived as a substitute for vaginal birth. However, cesarean delivery is associated with several disadvantages, including delayed initiation of breastfeeding, disrupted maternal-infant bonding, and increased risks in subsequent pregnancies. Moreover, the high cost of cesarean births places a considerable economic burden on national healthcare systems (3,4).

The childbirth process poses numerous risks to women during both the intrapartum and postpartum periods (5). Performing cesarean sections without valid medical indications can lead to significant adverse outcomes, including increased maternal and neonatal morbidity and mortality (1). Therefore, determining the mode of delivery is one of the critical decisions to be made during pregnancy.

Various factors influence a woman's choice regarding the mode of delivery. These may include previous birth experiences, recommendations from family and friends, concerns about neonatal health, preferences for specialized healthcare facilities, and participation in prenatal care visits (6). Women's perceptions of delivery methods are often shaped by their prior birth experiences.

The present study aims to explore the birth experiences of women with prior deliveries and to identify factors influencing their current preferences for mode of delivery when presenting at the clinic.

Materials and Methods

This descriptive cross-sectional survey was conducted at the Bahçelievler Memorial Hospital Obstetrics and Gynecology Outpatient Clinic between April 20, 2019 and January 20, 2020. Based on data from the first nine months of the previous year, the estimated number of women aged 15-49 years was approximately 960. Using Epi Info software and assuming a 95% confidence level, the minimum required sample size was calculated as 361 participants.

Two independent researchers informed eligible women (aged 15-49 years with a history of childbirth) about the study. Participants were asked about their previous modes of delivery, place of delivery, reasons for their delivery preferences, the decision-maker regarding the mode of birth, and experiences related to childbirth and to postpartum maternal and neonatal health. Socio-demographic information was also collected through a structured 31-item questionnaire.

Women who voluntarily agreed to participate and signed the informed consent form were interviewed face-to-face using the questionnaire during the study period. A total of 434 women completed the survey.

Ethical approval for the study was obtained from the Local Ethics Committee of Bezmialem Vakıf University on March 19, 2019 (approval no: 06/97).

Statistical Analysis

All statistical analyses were performed using IBM SPSS Statistics version 25.0. Continuous variables are presented as mean ± standard deviation, and categorical variables are

expressed as frequencies (n) and percentages (%). The chi-square test was used to compare categorical variables.

To identify predictors of the preferred mode of delivery in a future pregnancy, a multivariable logistic regression analysis was performed. A p-value <0.05 was considered statistically significant.

Results

The mean age of the participants was 32.87±6.76 years (range: 19-49), and 64.7% (n=281) had completed only primary education. While 363 (82.7%) of the participants were housewives, 21 (4.8%) were healthcare professionals. The frequency analysis of participants' demographic data is presented in Table 1.

At the time of the study, the mean number of vaginal births was 2.48±1.13 (range: 1-5), while the mean number of cesarean deliveries was 1.87±0.79 (range: 1-4). Frequency analysis of obstetric data is shown in Table 2. The most commonly reported reason for cesarean section was arrested labor (38.0%, n=113), followed by previous cesarean delivery (29.3%, n=87). Admission to the neonatal intensive care unit after delivery was higher among cesarean births (26.6%, n=50) than among vaginal births (16.7%, n=50).

Table 3 presents the association between responses to the question "If you were to become pregnant again, what delivery method would you prefer?" and various demographic and obstetric variables. Statistically significant associations were found between this question and the following variables: Current occupational sector (p=0.012), method of first delivery (p<0.001), method of most recent delivery (p<0.001), who made the decision regarding the first delivery (p<0.001), who made the decision regarding the most recent delivery (p=0.006), place of the first delivery (p<0.001), and, among those whose first delivery was vaginal, who decided on the delivery method (p=0.044).

To determine the factors influencing the preference for vaginal delivery in a future pregnancy, univariate and multivariate logistic regression analyses were performed (Table 4). In the initial step, each independent variable was analyzed separately, and those found to be significant were included in the multivariate model. The backward

Table 1. Socio-demographic data of the participants (n=434)

Demographic variables		N or $\bar{x}$ (min-max)	% or mean ± SD
Age		32.0 (19.0-49.0)	32.87±6.76
	19-30 years old	176	40.6
	31-40 years old	194	44.7
	Over 40 years old	64	14.7
Profession	Housewife	363	82.7
	Health worker	21	4.8
	Worker	36	8.2
	Other	19	4.3
Employment status	Yes	72	16.6
	No	362	83.4
Sector of employment (n=72)	Private sector	45	62.5
	State sector	27	37.5
Education status	Illiterate	39	9.0
	Primary education	281	64.7
	High school	64	14.7
	University	50	11.5
Smoking status	Yes	69	15.9
	No	365	84.1
Alcohol use status	Yes	20	4.6
	No	414	95.4
Do you have any chronic diseases?	No	379	87.3
	There is	55	12.7

N: Number, %: Frequency, $\bar{x}$: Median, Min: Minimum, Max: Maximum, Mean: Average, SD: Standard deviation

Table 2. The distribution of data regarding participants' obstetric characteristics

Demographic variables		N or $\bar{x}$ (min-max)	% or mean ± SD
How many births have you had so far?		2.0 (1.0-8.0)	2.47±1.08
	1 birth	75	17.3
	2 birth	171	39.4
	3 and above	188	43.3
Miscarriages (abortions) situation	No	337	77.6
	Yes	97	22.4
Curettage situation	No	354	81.6
	Yes	80	18.4
First delivery method	Vaginal	285	65.7
	Caesarean	149	34.3
Last delivery method	Vaginal	131	69.7
	Caesarean	57	30.3
Who decided the mode of birth? (first delivery)	Doctor	309	71.2
	Myself	125	28.8
Who decided the mode of birth? (last delivery)	Doctor	127	67.6
	Myself	61	32.4
Place of first birth	Private hospital/clinic	192	46.6
	State hospital	220	53.4
Place of last birth	Private hospital/clinic	67	37.6
	State hospital	111	62.4
How would you prefer to give birth if you were pregnant again?	Normal vaginal delivery	321	73.1
	Caesarean section	118	26.9

N: Number, %: Frequency, $\bar{x}$: Median, Min: Minimum, Max: Maximum, Mean: Average, SD: Standard deviation

likelihood-ratio method was used for variable selection in the multivariate model, producing a simpler, more reliable model by retaining only significant predictors with strong effects.

Only variables related to the first delivery were included in the model, as this experience was common to all participants and a defining event. The literature supports the notion that the first childbirth experience significantly shapes attitudes toward future delivery preferences and provides a more homogeneous foundation for analysis. Therefore, focusing solely on the first delivery improved the model's predictive power and validity.

In the univariate analyses, several variables were found to be significantly associated with a preference for vaginal birth in a future pregnancy, including vaginal birth in the first delivery [odds ratio (OR) =7.11; 95% confidence interval (CI): 4.46-11.35; p<0.001]; place of delivery (OR =2.52; 95% CI: 1.62-3.93; p<0.001); person deciding the delivery method (OR =2.84; 95% CI: 1.63-4.94; p<0.001); presence of chronic illness (OR =2.79; 95% CI: 1.56-4.99; p<0.001); regular use of medication (OR =2.35; 95% CI: 1.28-4.31; p=0.005); and employment sector (OR =2.96; 95% CI: 1.08-8.07; p=0.034). Occupation was also found to be significant (p=0.005); notably, healthcare professionals were less likely to prefer vaginal delivery (OR =0.21; p=0.001).

In the multivariate model, only three variables remained statistically significant predictors of a preference for vaginal delivery in a future pregnancy. Having a vaginal first delivery (OR =6.54; 95% CI: 3.94-10.85; p<0.001) and absence of chronic illness (OR =3.68; 95% CI: 1.86-7.28; p<0.001) were significantly associated with the outcome; giving birth in a public hospital (OR =1.61; 95% CI: 0.97-2.65; p=0.061) approached significance.

The overall model fit was acceptable (-2LL =401.74), and its explanatory power (Nagelkerke R^2) was 0.277.

Discussion

In our study, we explored women's perspectives on delivery methods by asking, "If you were to experience pregnancy again, which birthing method would you prefer?" The majority of participants expressed a preference for vaginal

Table 3. Comparison of "If you were to experience pregnancy again, which birthing method would you prefer?" with various variables

Variables	How would you prefer to give birth if you were pregnant again?			
	Group	Vaginal (n=316)	Caesarean section (n=118)	p
Which sector do you work in?	Private sector	33 (71.7)	12 (46.2)	0.031
	State sector	13 (28.3)	14 (53.8)	
Education status	Illiterate	32 (10.1)	7 (5.9)	0.109
	Primary education	208 (65.8)	73 (61.9)	
	High school	46 (14.6)	18 (15.3)	
	University	30 (9.5)	20 (16.9)	
Number of births	1 birth	58 (18.4)	17 (14.4)	0.012[a]
	2 birth	111 (35.1)	60 (50.8)	
	3 and above	147 (46.5)	41 (34.7)	
First mode of delivery	Vaginal	246 (77.8)	39 (33.1)	<0.001
	Caesarean section	70 (22.2)	79 (66.9)	
Last mode of delivery	Vaginal	113 (76.9)	18 (43.9)	<0.001
	Caesarean section	34 (23.1)	23 (56.1)	
Who decided the mode of first birth?	Doctor	209 (66.1)	100 (84.7)	<0.001
	Myself	107 (33.9)	18 (15.3)	
Who decided the mode of last birth?	Doctor	92 (62.6)	35 (85.4)	p<0.001
	Myself	55 (37.4)	6 (14.6)	
Place of first birth	Private hospital/clinic	119 (62)	73 (38)	<0.001
	State hospital	21 (91.3)	2 (8.7)	
Place of last birth	Private hospital/clinic	48 (35)	19 (46.3)	0.190
	State hospital	89 (65)	22 (53.7)	
If the first birth was vaginal, who decided on the mode of delivery?	Doctor	141 (57.3)	29 (74.4)	0.044
	Myself	105 (42.7)	10 (25.6)	

p<0.05, Pearson chi-square test

delivery. The most significant predictor of this preference was the mode of the first birth, with maternal health status also playing an influential role. However, the proportion of women preferring cesarean section was still higher than the rate recommended by the World Health Organization (WHO).

Over the past two decades, studies have documented increases, both globally and nationally (Turkey), in women's preference for cesarean section, even in the absence of medical indications (7). According to the WHO, the optimal cesarean delivery rate should be 10-15% (8). As of 2014, WHO reported cesarean section rates of 46% in Mexico, 32% in Germany, 33% in the United States, and 37% in Turkey (9). In our study, the vaginal delivery rate was 73.1%, contrasting with a study in China that reported 45.1% of births were vaginal and 54.9% were cesarean deliveries (10). Our cesarean delivery rates align with broader national trends.

Our analysis found that the mode of the first delivery significantly influenced preferences for the mode of subsequent deliveries. Notably, some women who had a cesarean for their first birth expressed a preference for vaginal delivery in a subsequent pregnancy. In line with our findings, Karabulutlu (11) reported that women with a previous vaginal birth were more likely to choose vaginal delivery again, whereas those with a previous cesarean were less likely to do so; this difference was statistically significant.

While cesarean delivery can be life-saving for both the mother and the infant when medically indicated, no evidence supports its benefits in the absence of such indications (12). Increased risks of allergies and childhood asthma among children born via cesarean section may influence the decision to opt for vaginal birth in subsequent pregnancies. Roman et al. (13) found that women with pre-existing illnesses or health issues during pregnancy

Table 4. Univariate and multivariate logistic regression results for the variables predicting preference for normal vaginal delivery

	Choosing a normal vaginal birth			
	University		Multivariate	
Variables	**OR (95% CI)**	**p**	**OR (95% CI)**	**p**
Constant	-	-	0.23	**<0.001**
Profession (Ref:Housewife)	-	**0.005**	-	-
Health worker	0.21 (0.08-0.53)	**0.001**	-	-
Employee	1.69 (0.68-4.19)	0.259	-	-
Other	0.75 (0.28-2.05)	0.585	-	-
Which sector do you work in? [Ref:(Government sector)]	2.96 (1.08-8.07)	**0.034**	-	-
Number of births	1.18 (0.96-1.45)	0.110	-	-
First mode of birth [Ref:(cesarean section)]	7.11 (4.46-11.35)	**<0.001**	6.54 (3.94-10.85)	**<0.001**
Who decided on the method of first birth? (Ref:Doctor)	2.84 (1.63-4.94)	**<0.001**	-	
Where did you give first birth? (Ref:Private hospital/clinic)	2.52 (1.62-3.93)	**<0.001**	1.61 (0.97-2.65)	0.061
Have a chronic disease? (Ref:Yes)	2.79 (1.56-4.99)	**<0.001**	3.68 (1.86-7.28)	**<0.001**
Do you use any medication regularly? (Ref:Yes)	2.35 (1.28-4.31)	**0.005**	-	-
Did you breastfeed your baby immediately after first birth? (Ref:No)	1.23 (0.56-2.69)	**0.592**	-	-

R^2: 0.277 -2 loglikelihood=401.74, Method: Backward LR, OR: Odds ratio, CI: Confidence interval

were more likely to deliver via cesarean section. However, in our study, women with chronic conditions were more likely to prefer vaginal birth, possibly because of associated benefits, such as higher breastfeeding rates, fewer neonatal intensive care unit admissions, and a lower risk of allergic conditions in infants.

A careful evaluation of delivery indications can help mitigate risks and guide women toward appropriate delivery choices. In a study by Mutlu and Yurtçu (14) lower cesarean rates were associated with higher vaginal delivery rates. In our study, most women delivered vaginally, and most reported that their physician made the decision regarding the delivery method. We believe this may reflect the impact of appropriate clinical indications and of national health policies favoring vaginal delivery.

Numerous studies have also shown that educational level, socio-economic status, and environmental factors play roles in determining delivery preferences (3,11,15,16). While Torloni et al. (17) reported higher cesarean rates among women with lower education levels, Gözükara and Eroğlu (3) found no significant association between education level and delivery method (18). In our study, we observed that women with lower education levels more frequently preferred vaginal delivery, whereas more educated women more often preferred cesarean section. A possible explanation is that a considerable proportion of the educated women in our sample were healthcare professionals.

Supporting this, studies in Turkey have shown that approximately half of obstetricians and 56% of healthcare workers consider cesarean delivery safer for the baby (19,20). Thus, the preference for cesarean delivery among educated women in our study may reflect a perception that it is safer.

Common medical indications for cesarean delivery in current practice include prior cesarean delivery, fetal distress, cephalopelvic disproportion, and malpresentation (especially breech) (2). In our study, the most frequently cited reason for cesarean was prolonged labor, followed by a previous cesarean, which is consistent with the literature.

Özkan et al. (15) showed that delivering in the private sector and a lack of information about delivery methods increased cesarean rates. Similarly, in our study, most women who underwent cesarean section did so based on physician-determined indications. Vaginal delivery was more common in public hospitals, while cesarean delivery rates were higher in the private sector. According to Turkey's 2015 health statistics, 70.5% of deliveries occurred in private hospitals, and the institutional cesarean rate was 53.1% (21). In contrast, in our study, more than half of the births took place in public hospitals, which may explain the higher vaginal delivery rate observed.

However, not all studies agree. Karabulutlu (11) found no significant relationship between delivery preference and place of birth.

Study Limitations

This study has several limitations. It is based on self-reported data, making it susceptible to recall bias. Moreover, no objective measurement tools were employed, and the study was conducted within a culturally and socio-economically homogeneous population, which may limit the generalizability of the findings.

Conclusion

The mode of a previous birth significantly influences women's preferences for the subsequent mode of delivery. Our findings indicate that women with a history of cesarean section often express a preference for vaginal delivery in future pregnancies. Experiences during and after prior births appear to influence this decision-making process. Additionally, women whose delivery mode was determined by their physician were more likely to undergo vaginal delivery.

It is essential that medical indications be carefully evaluated for each pregnancy, as the chosen delivery method can have a lasting impact on preferences for future births. Educating expectant mothers about delivery options during prenatal care is crucial. Developing evidence-based health policies that promote vaginal birth, while clearly communicating the risks associated with elective cesarean delivery, is equally important.

When planning the mode of delivery, women should receive balanced and individualized counseling on the potential risks and benefits of both vaginal birth and cesarean section. This approach helps prepare them for a variety of outcomes and strengthens their involvement in the decision-making process. Prenatal education programs and the active engagement of family physicians play a critical role in empowering pregnant women and providing the necessary guidance and support throughout the prenatal period.

Ethics

Ethics Committee Approval: Ethical committee approval was obtained from the Local Ethic Committee of Bezmialem Vakıf University on 19.03.2019 with the number 06/97 for the study.

Informed Consent: Women who voluntarily agreed to participate and signed the informed consent form were interviewed face-to-face using the questionnaire during the study period.

Footnotes

Authorship Contributions

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

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

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

References

1. Molina G, Weiser TG, Lipsitz SR, Esquivel MM, Uribe-Leitz T, Azad T, et al. Relationship between cesarean delivery rate and maternal and neonatal mortality. JAMA. 2015;314(21):2263-2270.
2. Betran AP, Torloni MR, Zhang JJ, Gülmezoglu AM; WHO working group on caesarean section. WHO statement on caesarean section rates. BJOG. 2016;123(5):667-670.
3. Gözükara F, Eroğlu K. Ilk doğumunu yapmış kadınların (primipar) doğum şekline yönelik tercihlerini etkileyen faktörler [Factors that effect the choices of primipars on the mode of delivery]. Hacettepe Üniversitesi Hemşirelik Fakültesi Dergisi. 2008;15(1):32-46. Turkish.
4. Sağlık Bakanlığı. Doğum ve sezaryen eylemi yönetim rehberi 2010. Erişim adresi: https://saglik.gov.tr/TR-11152/dogum-ve-sezaryen-eylemi-yonetim-rehberi---2010--42.html
5. Taşkın L. Obstetrics and women's health nursing. Obstetrics and Women's Health Nursing IX edition, 2009;9:275.
6. Mia MN, Islam MZ, Chowdhury MR, Razzaque A, Chin B, Rahman MS. Socio-demographic, health and institutional determinants of caesarean section among the poorest segment of the urban population: evidence from selected slums in Dhaka, Bangladesh. SSM Popul Health. 2019;8:100415.
7. Fobelets M, Beeckman K, Healy P, Grylka-Baeschlin S, Nicoletti J, Devane D, et al. Health economic analysis of a cluster-randomised trial (OptiBIRTH) designed to increase rates of vaginal birth after caesarean section. BJOG. 2019;126(8):1043-1051.
8. Eyi EGY, Mollamahmutoglu L. An analysis of the high cesarean section rates in Turkey by Robson classification. J Matern Fetal Neonatal Med. 2021;34(16):2682-2692.
9. World Health Organization. Table of contents and introduction. 2015. p. 92-6 World Health Statistics 2014.
10. Fan H, Gu H, You H, Xu X, Kou Y, Yang N. Social determinants of delivery mode in Jiangsu, China. BMC Pregnancy Childbirth. 2019;19(1):473.
11. Karabulutlu O. Factors affecting women's birth type preferences. Florence Nightingale J Nurs. 2013;20(3):210-218.
12. Cook JR, Jarvis S, Knight M, Dhanjal MK. Multiple repeat caesarean section in the UK: incidence and consequences to mother and child. A national, prospective, cohort study. BJOG. 2013;120(1):85-91.
13. Roman H, Blondel B, Bréart G, Goffinet F. Do risk factors for elective cesarean section differ from those of cesarean section during labor in low risk pregnancies? J Perinat Med. 2008;36(4):297-305.

14. Mutlu S, Yurtçu E. The effects of the social environment and socio-economic conditions of pregnant relatives of pregnant women in Karabük on the way of birth. Konuralp Medical Journal. 2020;12(2):296-300.

15. Özkan S, Aksakal F, Avcı E, Civil E, Tunca M. Kadınların doğum yöntemi tercihi ve ilişkili faktörler [Delivery methods: choices for women and related factors]. TJPH. 2013;11(2):59-71. Turkish.

16. Doganer G, Doganer A, Duran O, Tuna M, Tezcan S, Turan G. The determination of opinions on preferences of health sciences students on the mode of delivery. J Turk Soc Obstet Gynecol. 2013;10(1):31-36.

17. Torloni MR, Betrán AP, Montilla P, Scolaro E, Seuc A, Mazzoni A, et al. Do Italian women prefer cesarean section? Results from a survey on mode of delivery preferences. BMC Pregnancy Childbirth. 2013;13:78.

18. Akarsu RH, Mucuk S. Turkish women's opinions about cesarean delivery. Pak J Med Sci. 2014;30(6):1308-1313.

19. Arikan DC, Ozer A, Arikan I, Coskun A, Kiran H. Turkish obstetricians' personal preference for mode of delivery and attitude toward cesarean delivery on maternal request. Arch Gynecol Obstet. 2011;284(3):543-549.

20. Koken G, Cosar E, Sahin FK, Tolga Arioz D, Duman Z, Aral I. Attitudes towards mode of delivery and cesarean on demand in Turkey. Int J Gynaecol Obstet. 2007;99(3):233-235.

21. Turkish Ministry of Health. Health statistics yearbook. Available from: https://ekutuphane.saglik.gov.tr/Ekutuphane/kitaplar/2015-yili29pdf.pdf

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):227-234
DOI: 10.4274/BMB.galenos.2025.86547

Evaluation of Exhaled Carbon Monoxide Levels, Nicotine Dependence, and Psychosocial Factors in the Process of Smoking Cessation Treatment

Sigara Bırakma Tedavisi Sürecinde Ekspiryum Havası Karbonmonoksit Düzeyleri, Nikotin Bağımlılığı ve Psikososyal Faktörlerin Değerlendirilmesi

Hakan Gülmez, Kübra Erkek

İzmir Democracy University, Buca Seyfi Demirsoy Training and Research Hospital, Department of Family Medicine, İzmir, Turkey

Abstract

Objective: Tobacco use is the most important preventable cause of death and continues to be a major public health problem in our country and the world. Our aim in this study is to examine the changes in Fagerström nicotine dependance test (FNDT) score, depression-anxiety-stress (DASS-21) score, carbon monoxide (CO) levels in expiratory air and ppm level in control in our patients who applied to the smoking cessation clinic during their routine examination.

Method: This retrospective cross-sectional study included all patients with follow-up records at the smoking cessation clinic. The patient files were retrospectively reviewed for information on pack years of smoking, DAS score, and FNDT score obtained during tobacco dependence treatment monitoring system registration, as well as CO levels in expiratory air measured with the piCO Smokerlyzer during their application to our clinic. The SPSS-20 package program was used for statistical analysis. A value of $p<0.05$ was considered statistically significant.

Results: During follow-up, 76% (n=76) of participants quit smoking. Cessation rates did not differ by gender, but men had significantly higher pack-years and admission CO levels ($p<0.05$). In the Pearson correlation test, a significant positive correlation was found between FNDT and stress, depression, and anxiety scores. While there was a strong positive correlation between age and packs/years, a weak negative correlation was found between age and ppm/carboxyhemoglobin levels at admission. The mean pack-year value was found to be higher in the high dependency group.

Öz

Amaç: Tütün kullanımı önlenebilir ölüm nedenlerinden en önemlisi olup, ülkemizde ve dünyada halen büyük bir halk sağlığı problemi olmaya devam etmektedir. Bu çalışmada amacımız; sigara bırakma polikliniğine başvuran hastalarımızın rutin muayenesinde bakmış olduğumuz Fagerström nikotin bağımlılık testi (FNBT) skoru, depresyon-anksiyete-stres (DASS-21) skoru, ekspiryum havasındaki karbonmonoksit (CO) düzeyleri ve kontroldeki ppm düzeyi değişimini incelemektir.

Yöntem: Bu retrospektif kesitsel çalışmaya, sigara bırakma polikliniğine başvurup izlem kaydı bulunan tüm hastalar dahil edildi. Polikliniğimize başvuru esnasında sorgulanan sigara kullanım paket yılı, DAS skoru, tütün bağımlılığı tedavisi izlem sistemi kaydı sırasında elde edilen FNBT skoru, piCO Smokerlyzer ile ölçülen ekspiryum havasındaki CO düzeyleri geriye dönük olarak hasta dosyaları üzerinden taranmıştır. İstatistiksel analizde SPSS-20 paket programı kullanılmıştır. $p<0,05$ değeri istatistiksel olarak anlamlı kabul edilmiştir.

Bulgular: İzlem sürecinde katılımcıların %76'sı (n=76) sigarayı bıraktı. Bırakma oranı cinsiyetle anlamlı fark göstermedi; ancak erkeklerde paket yılı ve başvuru CO düzeyi anlamlı şekilde daha yüksekti ($p<0,05$). Yapılan Pearson korelasyon testinde FNBT ile stres, depresyon ve anksiyete skorları arasında anlamlı pozitif korelasyon bulunmuştur. Yaş ile paket/yıl arasında güçlü pozitif korelasyon olmasına karşın; yaş ile başvuru esnasındaki ppm/karboksihemoglobin düzeyleri

Address for Correspondence: Hakan Gülmez, İzmir Democracy University, Buca Seyfi Demirsoy Training and Research Hospital, Department of Family Medicine, İzmir, Turkey
E-mail: hakan.gulmez1@saglik.gov.tr **ORCID:** orcid.org/0000-0001-5467-3743
Received: 03.07.2025 **Accepted:** 26.11.2025 **Epub:** 01.12.2025 **Publication Date:** 24.06.2026

Cite this article as: Gülmez H, Erkek K. Evaluation of exhaled carbon monoxide levels, nicotine dependence, and psychosocial factors in the process of smoking cessation treatment. Bagcilar Med Bull. 2026;11(2):227-234

Abstract

Conclusion: CO is a toxic gas that reduces the amount of oxygen bound to tissues. Exhaled CO can be easily, quickly, and non-invasively used to monitor smoking cessation. Our findings suggest that exhaled CO levels may indirectly reflect nicotine dependence and psychological states. Large multicenter studies are needed to better identify factors influencing smoking cessation success.

Keywords: Carbonmonoxide, cytisine, smoking cessation

Öz

arasında zayıf negatif korelasyon saptanmıştır. Yüksek bağımlı grupta paket-yıl değeri ortalaması daha yüksek bulunmuştur.

Sonuç: CO dokulara bağlanan oksijen miktarını azaltan toksik bir gazdır. Sigara bırakma tedavisi, ekspiryumdaki CO ölçümüyle kolay, hızlı ve invaziv olmayan şekilde izlenebilir. Bulgularımız, ekspiryum havasındaki CO düzeyinin nikotin bağımlılığı ve psikolojik durum hakkında dolaylı bilgi verebileceğini göstermektedir. Sigara bırakma tedavisi başarısını etkileyen faktörlerin daha net bir şekilde ortaya konabilmesi için çok merkezli geniş katılımlı araştırmalara ihtiyaç vardır.

Anahtar kelimeler: Karbonmonoksit, sigara bırakma, sitizin

Introduction

Tobacco use is the leading cause of death and remains a major public health problem both in our country and globally. According to the World Health Organization, 1.3 billion people in the world use tobacco products, 80% of whom live in low- and middle-income groups. In Turkey, 150.4 billion cigarettes were consumed in 2024, marking the highest annual cigarette consumption in the Republic (1). Every year, approximately 7 million people die from tobacco use, and an additional 1.2 million die from exposure to tobacco smoke (2,3). More than 83,000 people lose their lives due to tobacco use each year (4).

To prevent the use of tobacco products, which are clearly known to pose a serious threat to public health, the World Health Organization has recommended the monitoring tobacco use; protecting people from tobacco smoke; offering help to quit smoking; warning the public about the harms of tobacco use; enforcing bans on tobacco advertising; raising taxes on tobacco products policy package to all member countries. Turkey was the first country to implement all these components and served as an example for other countries (2). Efforts to combat smoking, which remains the most common form of tobacco use, are ongoing.

Cigarettes contain more than 4.000 antigenic, carcinogenic, cytotoxic, and mutagenic substances (5). The burning of tobacco produces oxidizing chemicals (e.g., nitrosamines and polycyclic hydrocarbons), carcinogens, and toxic substances such as carbon monoxide (CO), metals, and particulate matter. These toxic substances cause various diseases, including cancer (6). CO is a colorless, tasteless, odorless, and non-irritating gas, and its affinity for hemoglobin is greater than that of oxygen. CO binds to hemoglobin, forming carboxyhemoglobin (COHb). The COHb that forms reduces oxygen delivery to the tissues. Normal COHb levels are 0.5-3% in adults, 4-12% in smokers, and 3-7% in newborns (7). The nicotine in cigarettes is addictive to smokers. Considering that addiction has psychological and social aspects, it is not always easy to overcome addiction, and supporting efforts to address addiction is important for both individual and public health.

The Fagerström nicotine dependence test (FNDT) is frequently used to determine the level of nicotine dependence and to plan treatment for individuals who smoke (8). CO levels in exhaled breath are also used as markers for diagnosis, treatment, and monitoring. Since it is directly related to smoking, it helps plan treatment objectively and increase individuals' compliance with treatment. Although no universally established cut-off value exists, studies suggest that exhaled CO levels above 6 ppm are highly indicative of active smoking (9).

Nicotine binds to nicotinic cholinergic receptors in the brain, modulating the activity of neurotransmitter systems such as dopamine, serotonin, and gamma-aminobutyric acid (GABA). This interaction particularly stimulates the reward mechanism within the mesolimbic dopaminergic pathway, forming the neurobiological basis of addiction. In addition, nicotinergic interactions stimulate the release of glutamate, GABA, and endorphins in various brain regions through both direct and indirect mechanisms, thereby mediating the wide range of physiological and behavioral effects of nicotine dependence beyond the reward system (10,11). It is also well established that the process of smoking cessation is influenced not only by biological factors but also by psychosocial determinants. The depression-anxiety-stress (DASS)-21 was employed in this study to simultaneously assess individuals' levels of depression, anxiety, and stress because of its multidimensional structure and well-documented validity and reliability

(12). By integrating psychosocial data with health records from the tobacco dependence treatment monitoring system (TUBATIS) database, the study aimed to provide a more comprehensive evaluation. Several studies have demonstrated an association between smoking behavior and psychological factors such as depression, anxiety, and stress (13-15).

Various methods, such as cognitive-behavioral therapy, nicotine replacement therapy, bupropion, and varenicline, are used in smoking cessation treatment, and the most recently introduced pharmacological agent in our country is cytisine. Cytisine, a plant-based alkaloid, binds to a4b2 nicotinic receptors, reducing nicotine withdrawal symptoms and the person's desire to smoke (16). The treatment protocol lasts approximately 25 days and is completed by gradually decreasing the drug dose.

The aim of this study is to examine the changes in FNDT scores, DASS-21 scores, and exhaled CO and COHb levels between the initial and follow-up assessments in patients who presented to the Smoking Cessation Outpatient Clinic of Buca Seyfi Demirsoy Training and Research Hospital, which has been providing services since March 2024.

Thus, the study aimed to evaluate the utility of CO and COHb measurements in combating nicotine dependence and to elucidate the bidirectional relationship between dependence levels (as measured by the FNDT) and psychiatric conditions.

Materials and Methods

This study employed a retrospective cross-sectional design. All patients who applied to the smoking cessation clinic, who were treated with cytisine, and who had a recorded follow-up examination were included in the study. Patients without follow-up records, with unknown smoking cessation status, or who were receiving treatment for depression or anxiety were excluded from the study.

The patient files were retrospectively reviewed to extract information on pack-years of smoking, DASS-21 and FNDT scores obtained during TUBATIS registration, as well as CO and COHb levels in exhaled air measured with the piCO Smokerlyzer at the time of their presentation to our clinic. Patients presenting between the opening of the clinic (01.03.2024) and 31.05.2025 were examined.

Ethics committee approval for the study was obtained from the Ethics Committee of Non-Interventional Research of Buca Seyfi Demirsoy Training and Research Hospital (approval number: 2025/457, date: 30 April 2025).

Statistical Analysis

Statistical analysis was performed using SPSS version 20. Descriptive distribution analysis, the chi-square test, and Student's t-test were used to analyze the data. A Pearson correlation test was performed to examine the relationships among variables. A value of $p<0.05$ was considered statistically significant.

Results

When retrospective patient records were examined, data from 100 individuals with control records who had been started on pharmacological treatment were obtained. All patients were started on a drug containing the active ingredient cytisine. 35% of the patients were male (n=35) and 65% were female (n=65). The median age was 42 years (range, 20-76) (Table 1). The average interval between the participants' initial visit and the follow-up assessment was 30 days. During the follow-up period, 76% (n=76) of participants quit smoking, while 24% (n=24) did not. While the smoking cessation rate in men was 65.7% (n=23) positive and 34.3% (n=12)

Table 1. Descriptive statistics

	Mean ± SD	Minimum	Maximum
Age	43.41±13.78	20	76
Pack-years of smoking	27.45±19.40	2	100
FNDT	6.49±2.48	1	12
Depression score	6.26±4.80	0	20
Anxiety score	6.54±4.33	0	20
Stress score	6.85±4.47	0	19
First application ppm	18.10±10.85	3	49
First application COHb	3.50±1.80	1.10	8.50
Control ppm	4.38±6.64	0	33
Control COHb	1.34±1.11	0.00	5.91

SD: Standard deviation

Table 2. The relationship between gender and smoking cessation status

			Smoking cessation status		Total
			Negative	Positive	
Gender	**Male**	**n**	12	23	35
		%	34.3	65.7	100
	Female	**n**	12	53	65
		%	18.5	81.5	100
Total		**n**	24	76	100
		%	24.0	76.0	100

$X^2=3.12$, $p=0.07$

negative, it was 81.5% (n=53) positive and 18.5% (n=12) negative in women (Table 2). There was no statistically significant association between smoking cessation status and gender (Figure 1). Pack-years of smoking, CO (ppm) measured in exhaled air at admission, and COHb levels were significantly higher in men (p<0.05; Table 3). There was a statistically significant difference in ppm change by gender between the ppm measurement at first admission and the ppm value detected at the control assessment; the ppm decrease was greater in men (p<0.05). A significant relationship was observed between the admission ppm value and the decrease in ppm (p<0.001). A significant relationship was observed between admission and control ppm/COHb levels (Table 4). The mean ppm for individuals who quit smoking 16.94 at admission and 1.90 at control; for those who could not quit smoking, it

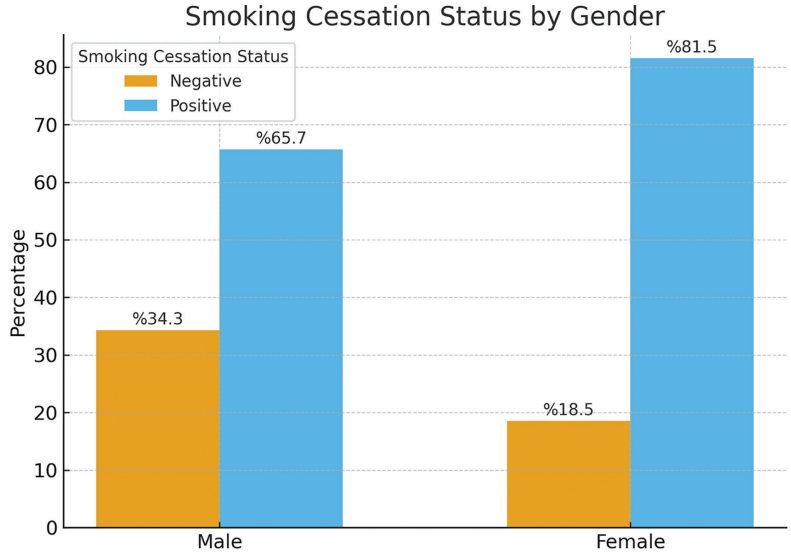


Figure 1. Smoking cessation status according to gender

Table 3. Comparison of clinical and psychological variables according to gender

	Gender	n	Mean	SD	p-value
Age	**M**	35	40.34	16.28	0.10
	F	65	45.06	12.04	
Pack-years of smoking	**M**	35	33.44	26.15	**0.02***
	F	65	24.22	13.74	
FNDT	**M**	34	6.71	2.53	0.54
	F	63	6.38	2.46	
Depression score	**M**	25	4.80	4.10	0.07
	F	56	6.91	4.97	
Anxiety score	**M**	25	5.76	3.90	0.28
	F	56	6.89	4.50	
Stress score	**M**	25	5.80	3.75	0.16
	F	56	7.32	4.71	
First application ppm	**M**	31	22.26	11.74	**0.01***
	F	56	15.80	9.69	
First application COHb	**M**	31	4.12	1.96	**0.02***
	F	56	3.16	1.62	
Control ppm	**M**	29	7.28	9.09	**0.003***
	F	55	2.85	4.25	
Control COHb	**M**	29	1.80	1.46	**0.004***
	F	55	1.09	0.78	

T-test, SD: Standard deviation, *: p<0.05, FNDT: Fagerström nicotine dependence test, COHb: Carboxyhemoglobin

was 22.53 at admission and 12.89 at control. There were statistically significant differences between the ppm levels at first admission and at control in both groups (p<0.001 for those who quit smoking; p=0.001 for those who could not quit). Using Pearson correlation analysis, significant positive correlations were found between FNDT and stress, depression, and anxiety scores (r=0.379, p<0.001; r=0.487, p<0.001; and r=0.386, p<0.001, respectively). While there was a strong positive correlation between age and pack-years, a weak negative correlation was found between age and ppm and COHb levels at admission (Table 5). When those with FNDT scores between 0-2 were classified as low dependency, those between 3-7 as moderate dependency, and those between 8-10 as highly dependent, 7 individuals had low, 48 individuals had moderate, and 42 individuals had high dependency. Since 3 patients had a TUBATIS record in the past, a new record could not be created, and their FNDT scores were not noted. According to the FNDT scores of the individuals, significant differences were found in some variables in low, moderate, and high dependency levels. No significant differences were found between the groups in terms of age, first application/control ppm, and COHb levels. On the other hand, a significant difference was found among the low-, moderate-, and high-dependency groups with respect to the cigarette pack-year variable (p=0.005). The mean pack-year value was higher in the high-dependency group. A significant difference was found between dependency groups for depression, anxiety, and stress scores (psychological variables) (p<0.05; Table 6).

Table 4. Comparison of baseline and follow-up exhaled CO values

	n	Mean ± SD	p-value
First application ppm	82	18.23±10.99	**<0.001***
Control ppm	82	4.45±6.71	
First application COHb	82	3.52±1.82	**0.001***
Control COHb	82	1.35±1.12	

Paired Samples t-test; SD: Standard deviation, *: p<0.001, COHb: Carboxyhemoglobin

Discussion

Expiratory CO measurement is used in many countries to objectively assess smoking because it is non-invasive, delivers rapid results, and is easy to perform. In studies conducted in Brazil in 2001 and 2007, expiratory CO values of smokers were significantly higher than those of non-smokers, and these values were reported to be useful for determining the probability that an individual is still smoking during smoking cessation treatment (17,18). Two recent studies in India have similarly found higher ppm concentrations among smokers (19,20). Similarly, in our study, the CO levels in the exhaled air of individuals who had quit smoking were significantly lower than those in smokers. A study conducted in 2018, similar to our study, showed a significant correlation between FNDT and pack-years of smoking (r=0.304, p<0.000). Although a positive correlation was found between FNDT and both age and ppm levels in the same study, no significant correlation was found in our study (21). In the study conducted by Kutlu et al. (22), the mean FNDT was 6.13±2.39 and the mean CO was 13.33±6.31 ppm; a statistically significant decrease was observed between the level at the time of application and that at the control, which is similar to our study. Babaoğlu et al. (23) showed that exhaled CO levels were higher in individuals with high addiction levels according to FNDT, consistent with our study. However, in the study, the FNDT scores of women (6±2.6) were significantly higher than those of men (5.2±2.2) (p=0.009). In contrast, no significant difference in nicotine addiction levels was found between genders in our study (23). In another study conducted by

Table 5. Relationship between age, pack-years of smoking, FNDT and psychological assessment results and CO measurement values

		Age	Pack-years of smoking	FNDT	Depression	Anxiety	Stress	First application ppm	First application COHb	Control ppm	Control COHb
Age	**r**	1	**0.57****	-0.0004	0.04	-0.04	-0.069	**-0.263***	**-0.271***	-0.083	-0.086
	p		**0.000****	0.97	0.76	0.75	0.540	**0.014**	**0.011**	0.452	0.439
Pack-years of smoking	**r**		1	**0.323***	0.102	0.048	0.019	0.059	0.036	0.039	0.017
	p			**0.001***	0.366	0.670	0.869	0.586	0.738	0.722	0.880
FNDT	**r**			1	**0.487****	**0.386****	**0.379****	0.155	0.120	0.148	0.116
	p				**0.000**	**0.000**	**0.001**	0.159	0.276	0.187	0.303

r: Pearson correlation coefficient; p: Significance level, **: Correlation is significant at the 0.001 level (2-tailed), *: Correlation is significant at the 0.05 level (2-tailed), CO: Carbon monoxide, FNDT: Fagerström nicotine dependence test, COHb: Carboxyhemoglobin

Table 6. Comparison of variables according to dependency level

		n	Mean	SD	p-value			Minimum	Maximum
					Low-moderate dependency	Moderate-high dependency	Low-high dependency		
Age	**Low dependency**	**7**	47.00	18.24	0.56	0.82	0.42	20	66
	Moderate dependency	**48**	43.35	14.87				20	76
	High dependency	**42**	42.71	11.87				20	66
	Total	**97**	43.34	13.80				20	76
Pack-years of smoking	**Low dependency**	**7**	28.43	21.0	0.31	**0.005***	0.58	2	50
	Moderate dependency	**48**	22.21	14.07				2	60
	High dependency	**42**	33.64	22.99				4	100
	Total	**97**	27.61	19.56				2	100
Depression	**Low dependency**	**6**	2.50	3.83	0.12	**0.008***	**0.001****	0	10
	Moderate dependency	**38**	5.21	3.93				0	16
	High dependency	**34**	8.03	4.85				1	20
	Total	**78**	6.23	4.63				0	20
Anxiety	**Low dependency**	**6**	3.33	2.66	0.1	**0.05***	**0.04***	1	8
	Moderate dependency	**38**	5.74	3.34				0	15
	High dependency	**34**	7.79	5.12				0	20
	Total	**78**	6.45	4.34				0	20
Stress	**Low dependency**	**6**	3.67	3.01	0.16	**0.03***	**0.03***	0	9
	Moderate dependency	**38**	6.08	3.99				0	15
	High dependency	**34**	8.32	4.77				0	19
	Total	**78**	6.87	4.48				0	19
First application ppm	**Low dependency**	**7**	17.29	12.66	0.82	0.15	0.60	5	35
	Moderate dependency	**39**	16.36	9.66				7	41
	High dependency	**38**	19.92	11.90				3	49
	Total	**84**	18.05	10.97				3	49
First application COHb	**Low dependency**	**7**	3.33	2.09	0.92	0.23	0.61	1.15	6.23
	Moderate dependency	**39**	3.26	1.57				1.25	7.20
	High dependency	**38**	3.76	2.01				1.10	8.50
	Total	**84**	3.49	1.82				1.10	8.50
Control ppm	**Low dependency**	**6**	3.50	3.02	0.98	0.32	0.66	1	9
	Moderate dependency	**38**	3.45	4.71				1	26
	High dependency	**37**	4.86	7.30				0	29
	Total	**81**	4.10	5.94				0	29
Control COHb	**Low dependency**	**6**	1.28	0.53	0.79	0.40	0.82	0.79	2.10
	Moderate dependency	**38**	1.19	0.77				0.43	4.79
	High dependency	**37**	1.39	1.25				0.00	5.67
	Total	**81**	1.29	1.0				0.00	5.67

T-test, SD: Standard deviation *: p<0.05, COHb: Carboxyhemoglobin

Çelepkolu (24), as in our study, no significant association was found between FNDT scores and age or gender. In the study conducted by Bohadana et al. (25), FNDT scores in men were reported to be significantly higher than those in women (6.44 and 5.99, respectively; p=0.018). In another study investigating the relationship between CO levels in exhaled air and FNDT, the average CO concentration was 13.2 ppm in smokers and 2.8 ppm in non-smokers. It was thought that ppm levels did not provide direct information about addiction severity, but were valuable for distinguishing smokers from non-smokers (26). In the literature, various scales have been employed to assess mental health status. In a study conducted among medical students, depression and anxiety scores, evaluated using the Beck depression inventory and the Beck anxiety inventory, respectively were higher in smokers than in non-smokers (13). Another study reported improvements in DASS-21 scores in patients with chronic obstructive pulmonary disease as their cigarette consumption decreased (27). Similarly, a study using the general health questionnaire-12 found that smokers exhibited higher levels of depression and anxiety than non-smokers. High nicotine dependence was associated with elevated depression and anxiety scores, consistent with the findings of our study (28).

The conflicting results in the literature regarding the relationship between FNDT scores and gender indicate that further studies are needed. The positive correlation between age and the pack-year value, which indicates cigarette consumption, is an expected consequence of longer smoking duration with increasing age. On the other hand, the smaller decrease in CO (ppm) levels in older age groups compared with younger age groups may be due to the adverse effects of long-term smoking on lung function. The significant decrease in the control CO levels among those who cannot quit smoking can be explained by increased awareness during the treatment process, increased motivational support, or decreased frequency of use.

Study Limitations

The limitations of our study include its retrospective design based on patient records, its single-center setting, and the lack of long-term follow-up for the patients. Since only patients who were initiated on cytisine and who attended regular check-ups were included, our study reports the success rate among cytisine-treated patients who attended follow-up. The fact that drug support can be provided free of charge may explain why individuals who apply to our smoking cessation clinic primarily seek drug treatment. In addition, non-attendance at follow-up appointments by individuals who have not been started on medication may be related to this situation. Three patients FNDT scores were not noted. Another limitation of our study is the higher proportion of female participants, which can be attributed to the inclusion only of individuals with follow-up records and to the generally greater tendency of women to seek healthcare services (29).

Conclusion

CO is a toxic gas that reduces the amount of oxygen bound to tissues. It binds to hemoglobin during smoking, forming the COHb complex and impairing tissue oxygenation. Smoking cessation treatment can be monitored easily, non-invasively, and rapidly by measuring CO levels in exhaled breath. Although the literature does not provide a clear-cut distinction between CO levels in smokers and non-smokers, the change in ppm from admission to follow-up indicates treatment response to both patient and physician and provides motivational support. In our study, significant relationships were found among addiction level, stress, anxiety, depression, and exhaled CO level; these findings suggest that exhaled CO measurement can provide indirect clues about nicotine addiction and individuals' psychological states. Multicenter studies with substantial participation are needed to better elucidate the factors affecting the success of smoking cessation treatment. In addition, cytisine requires a shorter treatment course than other therapies, which may increase patient compliance. However, since it has only recently been used in our country, a limited number of studies are available, and more comprehensive data on this subject are needed.

Ethics

Ethics Committee Approval: Ethics committee approval for the study was obtained from the Ethics Committee of Non-Interventional Research of Buca Seyfi Demirsoy Training and Research Hospital (approval number: 2025/457, date: 30 April 2025).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: H.G., K.E., Concept: H.G., K.E., Design: H.G., K.E., Data Collection or Processing: H.G., K.E., Analysis or Interpretation: H.G., K.E., Literature Search: H.G., K.E., Writing: H.G., K.E.

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

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

References

1. Tütün ve Tütün Mamulleri Daire Başkanlığı. Yıllar itibarıyla sigara iç satışı (1925–2024) [Internet]. Ankara: Tarım ve Orman Bakanlığı; [erişim tarihi:2025 Mayıs 25]. Erişim adresi:https://www.tarimorman.gov.tr/TADAB/Menu/22/Tutun-Ve-Tutun-Mamulleri-Daire-Baskanligi
2. T.C. Sağlık Bakanlığı Halk Sağlığı Genel Müdürlüğü. 2024-2028 Tütün kontrolü strateji belgesi ve eylem planı [Internet]. Ankara: T.C. Sağlık Bakanlığı; 2024 [erişim tarihi 2025 Mayıs 25]. Erişim adresi:https://hsgm.saglik.gov.tr/depo/Yayinlarimiz/Eylem_Planlari/2024-2028_Tutun_Kontrolu_Strateji_Belgesi_ve_Eylem_Plani.pdf
3. WHO report on the global tobacco epidemic, 2023 protect people from tobacco smoke fresh and alive Executive summary [Internet]. 2023. Available from:https://www.who.int/publications/book-orders.
4. Tütün Bağımlılığı. [Internet]. 2025 [erişim tarihi 2025 Mart 11]. Erişim adresi:https://www.yesilay.org.tr/tr/bagimlilik/tutun-bagimliligi
5. Karaözlü B, Karalezli A, editörler. Göğüs hastalıkları: Bölüm 14: Tütün kontrolü. İstanbul: Türkiye Solunum Araştırmaları Derneği; 2013.
6. Chang CM, Corey CG, Rostron BL, Apelberg BJ. Systematic review of cigar smoking and all cause and smoking related mortality. BMC Public Health. 2015;15:390.
7. T.C. Sağlık Bakanlığı. Sağlık çalışanları için karbonmonoksit zehirlenmeleri eğitim kitapçığı. [Internet]. 2016. Erişim adresi:https://hsgm.saglik.gov.tr/depo/birimler/cevre-sagligi-db/Dokumanlar/Sunumlar/CO_Saglik_Calisanlarina_Yonelik_Egit_Seti.pdf
8. Heatherton TF, Kozlowski LT, Frecker RC, Fagerström KO. The Fagerström test for nicotine dependence: a revision of the Fagerstrom tolerance questionnaire. Br J Addict. 1991;86(9):1119-1127.
9. Middleton ET, Morice AH. Breath carbon monoxide as an indication of smoking habit. Chest. 2000;117:758-763.
10. Benowitz NL. Neurobiology of nicotine addiction: implications for smoking cessation treatment. Am J Med. 2008;121(4 Suppl 1):S3-S10.
11. Duru Çetinkaya P, Arslan F. Türk Toraks Derneği tütün ürünleri ve nikotin bağımlılığında tanı ve tedavi uzlaşı raporu [Internet]. 2025. Erişim adresi:https://www.toraks.org.tr/site/sf/k/2025/04/735605b126d09929dfb6fc9c4e94f8321c47a54176b4b999c30c9c80e9ba02df.pdf
12. Yılmaz O, Boz H, Arslan A. The validity and reliability of depression stress and anxiety scale (DASS-21) Turkish short form. Research of Financial Economic and Social Studies. 2017;2(2):78-91.
13. Avcu B, Ulubaba Z, Ünlü UE, Ünlü N, Türkmenoğlu S, Unal MN, et al. Relationships between smoking and sleep quality, depression, and anxiety among medical students: a cross-sectional study in Türkiye. Turk J Tob Control. 2025;5(1):10-22.
14. Xiong H, Ma F, Tang D, Liu D. Correlations among nicotine dependence, health-related quality of life, and depression in current smokers: a cross-sectional study with a mediation model. Front Psychiatry. 2024;15:1455918.
15. Hu Z, Cui E, Chen B, Zhang M. Association between cigarette smoking and the risk of major psychiatric disorders: a systematic review and meta-analysis in depression, schizophrenia, and bipolar disorder. Front Med (Lausanne). 2025;12:1529191.
16. Kızıltaş Ö. New medical treatment for smoking cessation in Türkiye: cytisine. Journal of Dependence. 2025;26(2):274-276.
17. Santos UP, Gannam S, Abe JM, Esteves PB, Filho MF, Wakassa TB, et al. Emprego da determinação de monóxido de carbono no ar exalado para a detecção do consumo de tabaco *. J Pneumol. 2001;27(5):231-236. Portuguese.
18. Chatkin J, Fritscher L, de Abreu C, Cavalet-Blanco D, Chatkin G, Wagner M, et al. Exhaled carbon monoxide as a marker for evaluating smoking abstinence in a Brazilian population sample. Prim Care Respir J. 2007;16(1):36-40.
19. Vyshnavi SB, Shanthi M, Divya VC, Ganesh C, Gupta S, Jayavardhan K. Measurement of exhaled carbon monoxide levels in patients visiting to dental hospitals. Journal of Indian Academy of Oral Medicine and Radiology. 2024;36(3):292-296.
20. Shahbaz S, Nandedkar M. SA A, SS P, DN P. Evaluation of exhaled carbon monoxide levels in smokers and non-smokers-a comparative study. J Dental Sci. 2023;8(2):1-5.
21. Sugavanesh P, Pushpanjali K. Nicotine dependence, its risk indicators, and exhaled carbon monoxide levels among the smokers in Bengaluru, India. Indian J Community Med. 2018;43(3):220-223.
22. Kutlu R, Demirbaş N, Yazıcı T, Karaoğlu N. Evaluation of the effects of psychological addiction and smoking urge on the success of smoking cessation. Selcuk Med J. 2021;3(37):231-237.
23. Babaoğlu E, Karalezli A, Er M, Hasanoğlu HC, Öztuna D. Exhaled carbon monoxide is a marker of heavy nicotine dependence. Turk J Med Sci. 2016;46(6):1677-1681.
24. Çelepkolu T. The relationship between nicotine dependence level and age-gender among the smokers: Diyarbakir sample. Dicle Medical Journal. 2014;41(4):712-716.
25. Bohadana A, Nilsson F, Rasmussen T, Martinet Y. Gender differences in quit rates following smoking cessation with combination nicotine therapy: influence of baseline smoking behavior. Nicotine Tob Res. 2003;5(1):111-116.
26. Kapusta ND, Pietschnig J, Plener PL, Blüml V, Lesch OM, Walter H. Does breath carbon monoxide measure nicotine dependence? J Addict Dis. 2010;29(4):493-499. Erratum in: J Addict Dis. 2011;30(1):89.
27. Zarghami M, Taghizadeh F, Sharifpour A, Alipour A. Efficacy of smoking cessation on stress, anxiety, and depression in smokers with chronic obstructive pulmonary disease: a randomized controlled clinical trial. Addict Health. 2018;10(3):137-147.
28. Emre N, Topal K, Bozkurt N, Topaktas E. Mental health screening and increased risk for anxiety and depression among treatment-seeking smokers. Tob Induc Dis. 2014;12(1):20.
29. Redondo-Sendino Á, Guallar-Castillón P, Banegas JR, Rodríguez-Artalejo F. Gender differences in the utilization of health-care services among the older adult population of Spain. BMC Public Health. 2006;6:155.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):235-239
DOI: 10.4274/BMB.galenos.2026.52523

Clinical and Demographic Risk Profiles in the Early Neonatal Period: A Comparative Study of Turkish and Refugee Newborns

Yenidoğan Döneminde Klinik ve Demografik Risk Profili: Türk ve Mülteci Bebeklerin Karşılaştırmalı Değerlendirmesi

Yakup Karakurt, Emrah Can

University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Neonatal Intensive Care, İstanbul, Turkey

Abstract

Objective: This study explores how basic clinical and demographic factors relate to early neonatal sepsis in term newborns, with a focus on differences between Turkish and refugee populations.

Method: We conducted a retrospective analysis of 1,637 term newborns delivered at a tertiary public hospital in İstanbul between May 2023 and May 2025. Collected data included birth weight, Apgar scores, cord blood lactate levels, maternal nationality, and antenatal care status. Suspected early-onset sepsis was defined using clinical signs observed within the first 72 hours. We applied chi-square tests, correlation analysis, and multivariate logistic regression for statistical evaluation.

Results: Fewer refugee mothers received adequate antenatal care than Turkish mothers (58.4% vs. 74.9%, p<0.01). Suspected sepsis was also more common in refugee newborns (22.5% vs. 11.1%, p=0.03). Lower Apgar scores at one minute and elevated cord lactate levels emerged as independent predictors of early-onset sepsis. A moderate inverse relationship was also found between birth weight and lactate (r=-0.42, p<0.01).

Conclusion: Easily measurable clinical parameters, such as Apgar scores and cord blood lactate, can provide early indicators of sepsis risk. Differences in antenatal care rates may contribute to this risk, particularly among refugee populations. Tailored follow-up strategies could help bridge this gap.

Keywords: Apgar score, lactate, newborn, prenatal care, refugee health, sepsis

Öz

Amaç: Bu çalışma, term yenidoğanlarda erken neonatal sepsis ile ilişkili temel klinik ve demografik faktörleri incelemeyi ve Türk ile mülteci popülasyonlar arasındaki farkları ortaya koymayı amaçlamaktadır.

Yöntem: Mayıs 2023 ile Mayıs 2025 tarihleri arasında İstanbul'daki üçüncü basamak bir kamu hastanesinde doğan 1.637 term yenidoğan retrospektif olarak analiz edildi. Toplanan veriler arasında doğum ağırlığı, Apgar skorları, göbek kordonu laktat düzeyi, annenin milliyeti ve antenatal bakım durumu yer aldı. İlk 72 saat içinde gözlenen klinik bulgulara göre erken sepsis şüphesi tanımlandı. İstatistiksel analizde ki-kare testi, korelasyon analizi ve çok değişkenli lojistik regresyon kullanıldı.

Bulgular: Mülteci annelerin yeterli antenatal bakım alma oranı Türk annelere göre daha düşüktü (%58,4 vs. %74,9, p<0,01). Sepsis şüphesi mülteci yenidoğanlarda daha yüksekti (%22,5 vs. %11,1, p=0,03). Düşük 1. dakika Apgar skoru ve yüksek kordon laktat düzeyleri, erken sepsis için bağımsız belirleyiciler olarak öne çıktı. Ayrıca doğum ağırlığı ile laktat düzeyi arasında orta derecede negatif bir ilişki bulundu (r=-0,42, p<0,01).

Sonuç: Apgar skoru ve göbek kordonu laktat düzeyi gibi kolay ölçülebilen parametreler, erken sepsis riski hakkında önemli ipuçları sağlayabilir. Antenatal bakım farkları özellikle mülteci gruplarda bu riski artırabilir. Bu nedenle bireyselleştirilmiş takip stratejileri geliştirilmesi faydalı olacaktır.

Anahtar kelimeler: Apgar skoru, göçmen sağlığı, laktat, prenatal bakım, sepsis, yenidoğan

Address for Correspondence: Yakup Karakurt, MD, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Neonatal Intensive Care, İstanbul, Turkey
E-mail: ykp.karakurt@gmail.com **ORCID:** orcid.org/0009-0003-9092-9808
Received: 22.08.2025 **Accepted:** 24.02.2026 **Epub:** 06.03.2026 **Publication Date:** 24.06.2026

Cite this article as: Karakurt Y, Can E. Clinical and demographic risk profiles in the early neonatal period: a comparative study of Turkish and refugee newborns. Bagcilar Med Bull. 2026;11(2):235-239

Introduction

Neonatal sepsis, particularly when it arises within the first 72 hours after birth, continues to be a leading cause of illness and death in newborns across the globe (1). One of the main challenges with early-onset sepsis (EOS) is its vague, non-specific symptoms. Newborns often present with subtle signs, such as breathing difficulties, temperature instability, or feeding problems, which can hinder timely diagnosis (1).

Because there are no definitive early diagnostic tools, clinicians rely on simple and widely available indicators to assess risk. Parameters such as Apgar scores, birth weight, and umbilical cord blood gas values have long been used to evaluate a newborn's initial adaptation to life and physiological stress (2,3). In particular, cord blood lactate has received increasing attention as a potential marker of perinatal hypoxia and overall systemic strain. Elevated lactate levels, especially among infants with lower birth weights, have been linked to higher risks of early complications, including sepsis (4).

However, clinical markers alone do not provide a complete assessment. Broader social and environmental factors also shape neonatal health outcomes. In middle-income countries such as Turkey, refugee and migrant mothers often encounter significant barriers to accessing routine prenatal care. These gaps in care have been associated with increased rates of neonatal infections and poorer early health outcomes for newborns in these communities (5-7). To address such disparities, the World Health Organization (WHO) recommends a minimum of four antenatal visits during pregnancy to support maternal and newborn well-being (8).

This study was conducted at a large public hospital in İstanbul where both Turkish and refugee mothers receive care under shared institutional protocols. The main objective was to explore whether maternal demographic characteristics [such as nationality and antenatal care (ANC) attendance] and clinical indicators (such as the Apgar score and cord blood lactate) are associated with the risk of EOS in term neonates.

Materials and Methods

This retrospective analysis was conducted at University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, a public tertiary-referral center in İstanbul, Turkey. The study population included term neonates (gestational age ≥37 weeks) who were born between May 1, 2023, and May 31, 2025. Ethical approval for the study was obtained from the Institutional Review Board of University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital (approval no: 2025/08/19/082, date: 02.06.2025).

Eligible participants were singleton live births who had complete perinatal records and available umbilical cord blood gas data. Newborns with congenital anomalies, with incomplete data, or those transferred from other healthcare facilities were excluded from the analysis.

Data were extracted from the hospital's electronic medical system and included both maternal and neonatal variables. These variables included maternal age, nationality (classified as Turkish or refugee), number of ANC visits, mode of delivery, gestational age, birth weight, 1- and 5-minute Apgar scores, cord blood pH and lactate levels, and early neonatal outcomes. According to WHO guidelines, adequate ANC was defined as attending at least four prenatal visits (8). Maternal infection was standardized as a binary variable, defined as a physician-diagnosed bacterial or viral illness that required pharmacologic treatment during the current pregnancy, ensuring that subjective symptoms alone were not classified as infection.

The primary outcome was the incidence of suspected EOS within the first 72 hours. Sepsis was classified into two categories: Culture-proven sepsis, defined as a positive blood culture with clinical signs; and clinical sepsis, defined as the presence of at least two clinical signs (e.g., respiratory distress or temperature instability) accompanied by abnormal laboratory markers, such as elevated CRP or leukocytosis, necessitating empirical antibiotic treatment. In line with standard neonatal practice, empirical antibiotic treatment initiated within the first three days of life was also considered part of the diagnostic criteria (1).

Statistical Analysis

All statistical analyses were performed using SPSS version 28. Descriptive statistics were presented as means with standard deviations for continuous variables and as frequencies and percentages for categorical ones. Comparisons between groups were made using independent t-tests for continuous variables and chi-square tests for categorical variables. Pearson correlation coefficient was used to assess the relationship between birth weight and cord blood lactate levels. Variables with p-values under 0.10 in univariate analyses were included in a multivariate logistic regression model to determine independent predictors of suspected EOS. While the sample size (n=1.637) provided robust statistical power, interpretation of the multivariate model is limited by the retrospective single-center design

and reliance on clinical sepsis diagnoses, for which culture confirmation was not always available. A p-value less than 0.05 was considered statistically significant.

Results

Demographic Characteristics

A total of 1.637 term neonates were included in the final analysis. Of these, 1.026 (62.7%) were born to Turkish mothers, while 611 (37.3%) were born to refugee mothers. The average maternal age was 29.7±5.8 years in the Turkish group and 27.2±6.5 years in the refugee group, showing a statistically significant difference (p=0.01).

Turkish mothers were also significantly more likely to have received adequate ANC, with 74.9% meeting the threshold of four or more visits, compared to 58.4% among refugee mothers (χ^2=12.8, p<0.01; see Table 1). There were no statistically significant differences between the two groups regarding mode of delivery (χ^2=1.9, p=0.17) or sex distribution of the neonates (χ^2=0.8, p=0.36).

Sepsis Prevalence

Overall, 251 neonates (14.9%) were classified as having suspected EOS. The prevalence of suspected sepsis was notably higher among refugee newborns (22.5%) compared to Turkish newborns (11.1%), a difference that reached statistical significance (χ^2=9.1, p=0.03; see Figure 1 and Table 1).

Newborns with suspected sepsis had significantly lower 1-minute Apgar scores and higher average cord blood lactate levels than those without sepsis (both p<0.01).

The mean birth weight across the sample was 3.160±340 grams, and the mean cord blood lactate concentration was 2.9±1.2 mmol/L. There was a moderate but significant inverse correlation between birth weight and lactate level (r=-0.42, p<0.01; see Figure 2). Neonates with 1-minute Apgar scores below 7 had higher lactate levels and were more frequently diagnosed with suspected sepsis (p<0.01).

In multivariate logistic regression (Table 2), two factors emerged as independent predictors of suspected EOS: a low 1-minute Apgar score [odds ratio (OR) =2.9; 95% confidence interval (CI): 1.6-5.3; p<0.01] and an elevated cord blood lactate level above 3.0 mmol/L (OR=2.2; 95% CI: 1.1-4.4; p=0.02). Although maternal nationality and ANC were associated with sepsis risk in univariate analysis, they were not retained as significant in the final model.

All neonates included in the study survived to hospital discharge. No in-hospital mortality was observed during the study period.

Discussion

In this study, we explored both clinical and demographic factors that may predict EOS in term infants born at a public tertiary hospital. Our findings show that low 1-minute Apgar scores and elevated cord blood lactate levels are strongly associated with an increased likelihood of suspected sepsis. Additionally, refugee status and insufficient ANC were linked to higher rates of suspected sepsis in the univariate analysis.

EOS continues to be a leading cause of illness and death among newborns, particularly within the first 72 hours of life. However, diagnosing it early remains difficult due to the non-specific nature of its clinical signs and the limitations of rapid confirmatory testing. As a result, healthcare providers often turn to readily available indicators to guide early intervention strategies (1).

Among these indicators, the 1-minute Apgar score is one of the most universally used measures of a newborn's condition. Previous research has associated low Apgar scores with adverse outcomes, including sepsis, particularly when scores are below 7 (2,3). Our results are consistent with that pattern: newborns with lower Apgar scores had nearly threefold higher odds of being treated for suspected EOS.

Cord blood lactate has emerged as another valuable marker in recent years. In our dataset, we observed that higher lactate levels were more common in neonates with lower birth weight. This supports the idea that smaller neonates may have reduced oxygen reserves or greater metabolic

Table 1. Demographic and clinical characteristics by maternal nationality

Characteristic	Turkish	Refugee	p-value
Maternal age (years)	29.7±5.8	27.2±6.5	0.01
Adequate antenatal care (%)	74.9	58.4	<0.01
Sepsis suspicion (%)	11.1	22.5	0.03
Birth weight (g)	3160±340	3080±360	0.06
Lactate (mmol/L)	2.7±1.1	3.1±1.2	0.04

stress at birth (4). Notably, elevated lactate levels remained a significant predictor of suspected sepsis in our multivariate model, consistent with earlier studies (5).

Although maternal nationality did not show a significant independent effect in the multivariate analysis, refugee newborns had notably higher rates of suspected sepsis. These infants were also more likely to have been born to mothers who did not receive adequate prenatal care. This finding echoes national data from Turkey, which indicate that refugee women often encounter systemic barriers to accessing antenatal services, including language challenges, limited health literacy, and unfamiliarity with the healthcare system (6,7).

The World Health Organization recommends at least four antenatal visits to support maternal and neonatal well-being (8). In our cohort, fewer than 60% of refugee mothers reached this benchmark, compared to nearly 75% of Turkish mothers.

This gap is significant. Recent studies from Turkey have highlighted that limited prenatal care among refugee populations is associated with a higher incidence of neonatal complications, including infections and preterm births (9,10). These disparities persist even when medical protocols are standardized, which underscores the importance of culturally responsive care models and targeted support at both community and system levels.

One limitation of our analysis is that detailed maternal health data, such as anemia, infections during pregnancy, hypertensive disorders, or premature rupture of membranes, were not consistently recorded in the hospital's retrospective dataset. As a result, we could not fully adjust for these potential confounders, which may have influenced the associations observed.

Study Limitations

This study has several important limitations. First, as a retrospective analysis, it is inherently subject to documentation bias and cannot establish causal relationships. Second, the identification of EOS was based on clinical judgment and laboratory findings rather than on confirmed blood cultures in every case. While this approach reflects routine clinical practice in many public hospitals, it could have led to an overestimation of the incidence of sepsis.

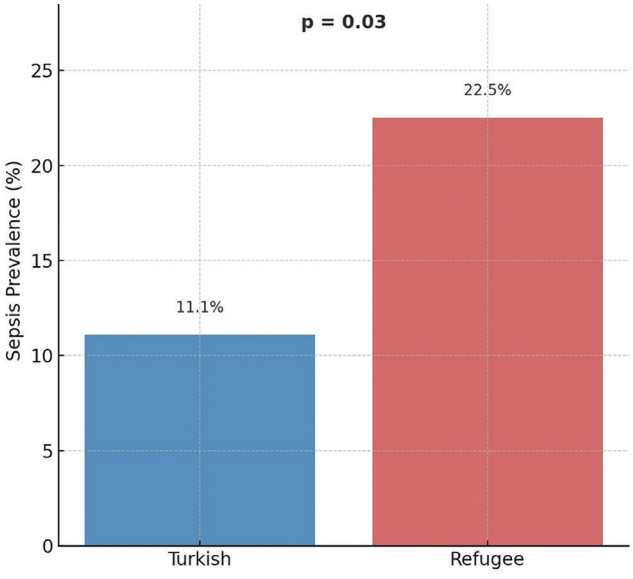


Figure 1. Sepsis prevalence by nationality (chi square test, χ^2=9.1, p=0.03, n=1.637)

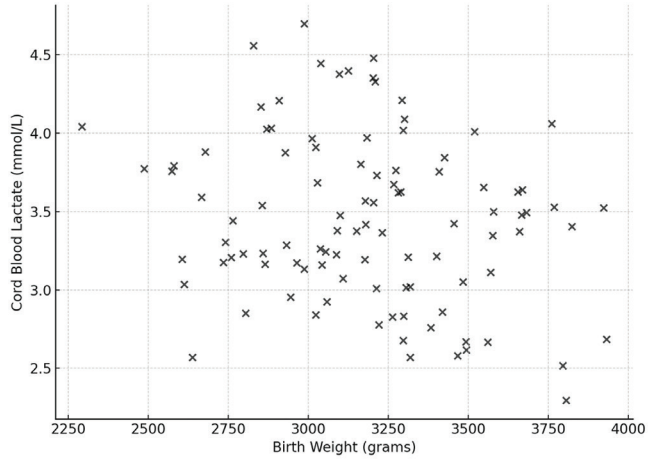


Figure 2. Relationship between birth weight and cord blood lactate (Pearson correlation, r=-0.42, p<0.01, n=1.637)

Table 2. Logistic regression analysis for predictors of sepsis

Variable	Adjusted OR	95% CI	p-value
Low 1-min Apgar score (<7)	2.9	1.6-5.3	<0.01
Elevated lactate (>3.0 mmol/L)	2.2	1.1-4.4	0.02
Maternal nationality (ref: Turkish)	1.5	0.9-2.6	0.09
Inadequate antenatal care	1.4	0.8-2.5	0.11

OR: Odds ratio, CI: Confidence interval

Third, although our sample size was relatively large, the data were collected from a single-center. This limits the generalizability of the findings, particularly to settings with different population dynamics or healthcare infrastructure.

Additionally, some potentially influential maternal health factors, such as anemia, infections during pregnancy, hypertensive disorders, and premature rupture of membranes were not consistently documented in the hospital records and were therefore excluded from the analysis. Prospective studies that include these variables could offer a more complete picture of the maternal contributions to neonatal sepsis risk.

Conclusion

Our results highlight the value of using simple, readily available indicators such as Apgar scores and cord blood lactate levels for early risk assessment in term newborns. When interpreted alongside maternal sociodemographic factors, these tools may help clinicians better identify infants at higher risk of EOS.

The higher rate of suspected sepsis among refugee newborns emphasizes the need for targeted prenatal interventions, particularly those that improve access to culturally sensitive and consistent ANC.

Importantly, all neonates in this study survived to discharge, suggesting that timely recognition and adherence to standardized management protocols may contribute to favorable short-term outcomes.

Going forward, public health strategies that account for both medical and social risk factors will be essential for improving neonatal outcomes, especially in busy public hospital settings serving diverse populations.

Ethics

Ethics Committee Approval: Ethical approval for the study was obtained from the Institutional Review Board of University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital (approval no: 2025/08/19/082, date: 02.06.2025).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Concept: Y.K., E.C., Design: Y.K., E.C., Data Collection or Processing: Y.K., E.C., Analysis or Interpretation: Y.K., E.C., Literature Search: Y.K., E.C., Writing: Y.K., E.C.

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

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

References

1. Du LZ. Early diagnosis and management of neonatal sepsis: a perspective. World J Pediatr. 2024;20(4):303-306.
2. Ludorf KL, Benjamin RH, Canfield M, Shumate C, Findley TO, Johnson A, et al. Low Apgar score and risk of neonatal mortality among infants with birth defects. Am J Perinatol. 2025;42(8):1024-1034.
3. Badmus OM, Adenaya OR, Aderinwale OA, Ewuoso BO, Awolaja BS, Ade-Onojobi AO. Umbilical arterial blood lactate as predictor of early neonatal outcome and evaluation of intrapartum asphyxia. J Taibah Univ Med Sci. 2024;19(5):911-918.
4. Malin GL, Morris RK, Khan KS. Strength of association between umbilical cord pH and perinatal and long term outcomes: systematic review and meta-analysis. BMJ. 2010;340:c1471.
5. Nguyen HB, Rivers EP, Knoblich BP, Jacobsen G, Muzzin A, Ressler JA, et al. Early lactate clearance is associated with improved outcome in severe sepsis and septic shock. Crit Care Med. 2004;32(8):1637-1642.
6. Taşkıran D, Ay O. A five-year retrospective evaluation of perinatal outcomes in Syrian refugee and Turkish citizen pregnant women and their newborns at a tertiary hospital. BMC Pregnancy Childbirth. 2025;25(1):1175.
7. Özgürel Bozkurt A, Usal Tarhan N, Özkaya E. The pregnancy and newborn outcomes of Syrian refugees and Turkish women in a tertiary center, in İstanbul, Türkiye. Zeynep Kamil Med J. 2022;53(3):116-121.
8. World Health Organization. WHO recommendations on antenatal care for a positive pregnancy experience. Geneva: World Health Organization; 2016.
9. Benova L, Tunçalp Ö, Moran AC, Campbell OMR. Not just a number: examining coverage and content of antenatal care in low-income and middle-income countries. BMJ Glob Health. 2018;3(2):e000779.
10. Haddrill R, Jones GL, Mitchell CA, Anumba DO. Understanding delayed access to antenatal care: a qualitative interview study. BMC Pregnancy Childbirth. 2014;14:207.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):240-245
DOI: 10.4274/BMB.galenos.2026.96720

Reliability and Consistency of ChatGPT-4o's Responses to Frequently Asked Questions About Hypertension

ChatGPT-4o'nun Hipertansiyon Hakkında Sıkça Sorulan Sorulara Verdiği Yanıtların Güvenilirliği ve Tutarlılığı

Ekrem Bilal Karaayvaz, Berk Batuhan Bayraktar, Göksel Güz

İstanbul University, İstanbul Faculty of Medicine, Department of Cardiology, İstanbul, Turkey

Abstract

Objective: To evaluate the quality, validity and reliability of the answers given by ChatGPT-4o to frequently asked questions about systemic arterial hypertension (HT), which directly concerns public health.

Method: In this study, 30 frequently asked questions about HT from health forums and hospital websites were compiled and divided into four categories: General HT questions (n=10), treatment-related questions (n=10), specific questions (n=5), and questions based on the 2024 European Society of Cardiology (ESC) guidelines (n=5). Questions were asked in Turkish to ChatGPT-4o and the answers were evaluated by three cardiologists using the global quality scale (GQS). The consistency of the answers was tested by repeating the same questions on different days.

Results: 56.7% of the responses were rated as high quality and comprehensive answers (GQS 5), 40% as largely comprehensive and good quality answers (GQS 4) and 3.3% as moderate quality (GQS 3). 80% of the questions based on the ESC guidelines received a GQS 5 score and demonstrated 100% reproducibility.

Conclusion: ChatGPT-4o can provide reliable, accurate, and repeatable answers to frequently asked questions about HT. However, it should be noted that users should support this information with professional medical advice.

Keywords: Artificial intelligence, ChatGPT-4o, systemic hypertension

Öz

Amaç: Toplum sağlığını doğrudan ilgilendiren sistemik arteriyel hipertansiyon (HT) hakkında sıkça sorulan sorulara ChatGPT-4o'nun verdiği yanıtların kalite, geçerlilik ve güvenilirliğini değerlendirmektir.

Yöntem: Bu çalışmada sağlık forumları ve hastane web sitelerinde yer alan HT ile ilgili 30 sık sorulan soru derlenmiş ve dört kategoriye ayrılmıştır: Genel HT soruları (n=10), tedaviyle ilgili sorular (n=10), spesifik sorular (n=5) ve 2024 Avrupa Kardiyoloji Derneği (ESC) kılavuzuna dayalı sorular (n=5). Sorular ChatGPT-4o'ya Türkçe yöneltilmiş ve verilen yanıtlar üç kardiyoloji uzmanı tarafından global kalite ölçeği (GQS) ile değerlendirilmiştir. Yanıtların tutarlılığı, aynı soruların farklı günlerde tekrarlanarak sorulmasıyla test edilmiştir.

Bulgular: Yanıtların %56,7'si yüksek kaliteli ve kapsamlı yanıtlar (GQS 5), %40'ı büyük oranda kapsamlı ve iyi kalitede yanıtlar (GQS 4) ve %3,3'ü ise orta düzeyde kalite (GQS 3) olarak değerlendirilmiştir. ESC kılavuzuna dayalı soruların %80'i GQS 5 puanı almış ve %100 tekrarlanabilirlik göstermiştir.

Sonuç: ChatGPT-4o, HT ile ilgili sık sorulan sorulara güvenilir, doğru ve tekrarlanabilir yanıtlar verebilmektedir. Ancak kullanıcıların bu bilgileri profesyonel tıbbi danışmanlıkla desteklemesi gerektiği unutulmamalıdır.

Anahtar kelimeler: ChatGPT-4o, sistemik arteriyel hipertansiyon, yapay zeka

Address for Correspondence: Lec, Ekrem Bilal Karaayvaz, MD, İstanbul University, İstanbul Faculty of Medicine, Department of Cardiology, İstanbul, Turkey
E-mail: ekrembilal@gmail.com **ORCID:** orcid.org/0000-0002-0105-6167
Received: 26.09.2025 **Accepted:** 04.04.2026 **Epub:** 15.04.2026 **Publication Date:** 24.06.2026

Cite this article as: Karaayvaz EB, Bayraktar BB, Güz G. Reliability and consistency of ChatGPT-4o's responses to frequently asked questions about hypertension. Bagcilar Med Bull. 2026;11(2):240-245

Introduction

ChatGPT is an artificial intelligence (AI) language model developed by OpenAI and designed to produce texts that can mimic human conversations (1). This technology has recently become a platform for people seeking information on a wide range of topics. As with other curiosities, people seek to use ChatGPT as a source of health information. Having information across a wide spectrum, from symptoms and diagnosis to treatment options, is particularly appealing to people. It is not surprising that ChatGPT is preferred by people due to its proactive nature and its potential to provide them with additional information based on the information it obtains by offering them a wide range of options. The popularity of ChatGPT and people's desire to use this AI to obtain information on any subject, receive suggestions, and even make decisions raise the question of how reliable the answers given by this AI language model are (2-4). People who want to use this platform should be able to obtain accurate health information and assess how accurate, evidence-based, and up-to-date the information is. Some studies have examined ChatGPT's validity and reliability across diverse topics within various health domains (5-8).

Systemic arterial hypertension (HT) is a major health problem affecting more than one billion adults worldwide (9). It has been shown that uncontrolled HT causes 9.4 million deaths and 212 million healthy life years lost each year (10). HT, a public health problem, is a health term frequently searched by people on internet search engines to obtain information about both diagnosis and treatment and follow-up (11). In this study, we aimed to investigate the quality, validity, and reliability of ChatGPT-4o's responses to HT-related questions, which are of particular concern to public health.

Materials and Methods

This study was conducted between March 18, 2025, and April 15, 2025. Since no patient data were used, informed consent was not required and therefore was not obtained.

Ethical approval for the study was obtained from the Institutional Ethics Committee of Medicana International İstanbul Hospital (decision number: 2025/2031; date: 24.09.2025).

The questions used in the study were compiled from frequently asked queries about HT on health forums and hospital websites. Questions for advertising purposes, those requiring personal responses, and repetitive questions were excluded from the study. Only questions written in Turkish were considered; a total of 30 were evaluated. The questions were divided into four groups: General questions about HT (10); questions about HT treatment (10); specific questions about HT (5); and questions related to the 2024 European Society of Cardiology (ESC) HT guidelines (12). The guidelines were reviewed, and 5 questions about HT were prepared. All questions are presented in Table 1.

Questions were directed to ChatGPT-4o in Turkish. The responses provided by ChatGPT-4o were evaluated by two cardiologists, each with at least 10 years of experience, and they were unaware of the score prior to the study. If the evaluations of the two physicians were the same, this score was recorded directly. If there was a difference between the evaluations, a third experienced cardiologist also performed an evaluation, and the final score was calculated as the arithmetic mean of the three scores. The consistency of ChatGPT's responses was tested by asking each question twice, on different days. Each question was submitted to ChatGPT-4o twice, with an interval of at least 7 days between queries. This interval was chosen to minimize potential short-term contextual or memory-related effects and to better assess the reproducibility of responses over time. The same ChatGPT account was used for all queries, ensuring consistency in model access. Each question was submitted in a new independent session, with chat history cleared prior to each query, so no prior prompts or responses were visible to the model. This approach was deliberately adopted to reduce contextual carryover effects and to ensure that each response was generated independently.

The quality and reliability of ChatGPT-4o responses were assessed using the global quality scale (GQS), developed to evaluate the accuracy and adequacy of medical content. GQS is a 1-to-5 scoring system used to assess the quality and reliability of written medical content.

According to scoring:

- GQS 1: Low quality, poorly organized, missing much of the essential information, and useless content for the patient.
- GQS 2: Overall poor structure and content, some information available but missing important topics, of little benefit to the patient.
- GQS 3: Moderate quality, some important information is adequately covered but others are inadequate, providing moderate benefit to the patient.

Table 1. Questions asked to ChatGPT-4o and answer scores

General questions about hypertension	**GQS**
1. What is hypertension?	4
2. What are the symptoms of hypertension?	5
3. How is hypertension diagnosed?	4
4. What are the causes of hypertension?	4
5. Are home blood pressure measurements from the wrist reliable?	5
6. Hypertension is more common in which age groups?	5
7. Is hypertension hereditary? Does my risk increase if it runs in my family?	4
8. Which organs are damaged by hypertension?	5
9. What is the relationship between hypertension and stress?	5
10. What kind of health problems can hypertension cause if left untreated?	5
Questions about hypertension treatment	
1. How is hypertension treated?	4
2. Does garlic and lemon juice lower blood pressure?	5
3. Are hypertension medications used for life?	5
4. How can blood pressure be lowered without medication?	4
5. How should hypertension patients eat?	5
6. How does salt consumption affect hypertension?	5
7. What lifestyle changes are recommended for hypertension?	4
8. How does exercise affect blood pressure?	4
9. Are alternative medicine methods effective in the treatment of hypertension?	3
10. Can hypertension be treated with surgery?	5
Specific questions about hypertension	
1. How is hypertension treated during pregnancy?	4
2. Does hypertension increase the risk of heart attack and stroke?	4
3. Does hypertension affect sexual life?	5
4. How alcohol and smoking affect hypertension?	4
5. Does blood pressure rise when visiting the doctor?	5
Hypertension questions based on the ESC guidelines	
1. How is hypertension diagnosed? What blood pressure levels are considered hypertension according to ESC guidelines?	5
2. What are the criteria for initiating antihypertensive treatment according to ESC guidelines?	5
3. How are treatment initiation thresholds determined based on blood pressure level, age and comorbidities?	5
4. According to ESC guidelines, which agents should be included in first-line drug therapy for hypertension? In what situations is combination therapy recommended?	4
5. According to ESC guidelines, what is the role of lifestyle changes in hypertension management? In which cases are measures such as diet, exercise, and salt restriction considered sufficient?	5

GQS: Global quality scale, ESC: European Society of Cardiology

• GQS 4: Good quality and structure, most important information is included, some omissions but useful for the patient.

• GQS 5: High quality and well organized, all important topics are covered comprehensively, extremely useful for the patient.

Statistical Analysis

All analyses were performed using the Statistical Package for the Social Sciences (SPSS) version 22.0 (SPSS Inc., Chicago, USA). Subcategory GQS scores are presented as percentages, and mean scores and repeatability rates as figures.

Interobserver reliability among the cardiologists evaluating the ChatGPT-4o responses was assessed using the intraclass correlation coefficient (ICC). Given that the GQS scores are ordinal and that more than two raters were involved, a two-way random-effects model for absolute agreement was applied. ICC values were interpreted according to commonly accepted criteria, with values below 0.50 indicating poor agreement, values of 0.50-0.75 indicating moderate agreement, values of 0.75-

0.90 indicating good agreement, and values above 0.90 indicating excellent agreement. Statistical analyses were performed using SPSS version 22.0 (SPSS Inc., Chicago, USA).

Results

In total, ChatGPT-4o answered 17 out of 30 questions (56.7%) comprehensively, providing high-quality, well-structured answers and earning a GQS score of 5. Twelve questions (40%) received a GQS score of 4, with answers that were largely comprehensive and of good quality, whereas only one question (3.3%) received a GQS score of 3, with answers of moderate quality and usefulness. The mean GQS score for the answers is shown in Figure 1.

6 (60%) of the responses to general HT questions were rated as high quality and comprehensive, receiving a GQS score of 5. Four (40%) responses received GQS scores of 4.

While 5 (50%) of the treatment-related questions received a GQS score of 5, 4 (40%) received a GQS score of 4. Only 1 (10%) question — "Are alternative medicine methods effective in the treatment of HT?"— received a GQS score of 3 because its content was insufficiently comprehensive and provided limited benefit.

In the specific questions category, 2 questions (40%) received a GQS 5 score, while the remaining 3 questions (60%) were rated as GQS 4.

Of ChatGPT-4o's responses to questions derived from the ESC guidelines, 4 (80%) were scored as GQS 5, indicating the highest quality and comprehensiveness, while 1 response was scored as GQS 4.

The reproducibility rates of ChatGPT-4o responses are presented in Figure 2. A reproducibility rate of 80% was observed for general and specific questions on HT. This rate was highest for questions based on ESC guidelines at 100%, while the lowest reproducibility rate was observed for questions on HT treatment at 70%.

Interobserver agreement analysis demonstrated good-to-excellent consistency among the cardiologists evaluating ChatGPT-4o responses. The calculated ICC was 0.82 (95% confidence interval: 0.76-0.87), indicating a high degree of agreement between expert raters and supporting the reliability of the scoring methodology employed in this study.

Discussion

Since its launch in November 2022, ChatGPT has quickly become a popular AI platform that both patients and clinicians use to obtain information and facilitate decision-making in diagnosis and treatment processes (13). In this study, in which we investigated the quality and validity of ChatGPT's answers to questions about HT, we observed that ChatGPT provided high-quality, structured answers to more than half of the questions (56.7%) and good-quality answers to 40% of the questions. Half of the answers to treatment-related questions were high-quality, comprehensive, and extremely useful for patients, and nearly half (40%) were determined to be good-quality and useful; these answers had a GQS score of 4. According to the results of our study, ChatGPT's answers to general HT and treatment questions were high-quality, but it is not sufficient to prove that these have clinical accuracy.

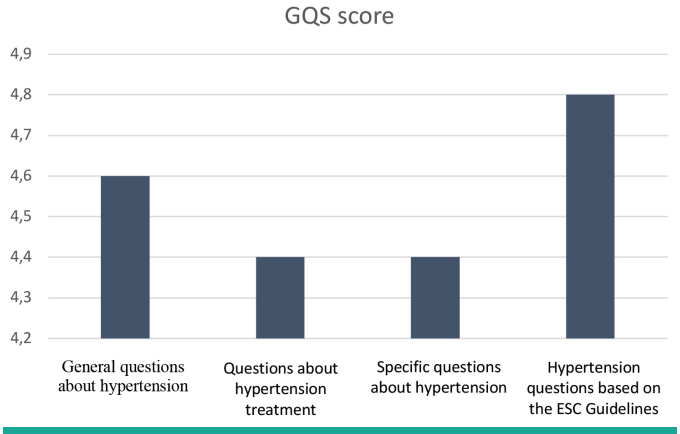


Figure 1. Average GQS scores of question answers
GQS: Global quality scale, ESC: European Society of Cardiology

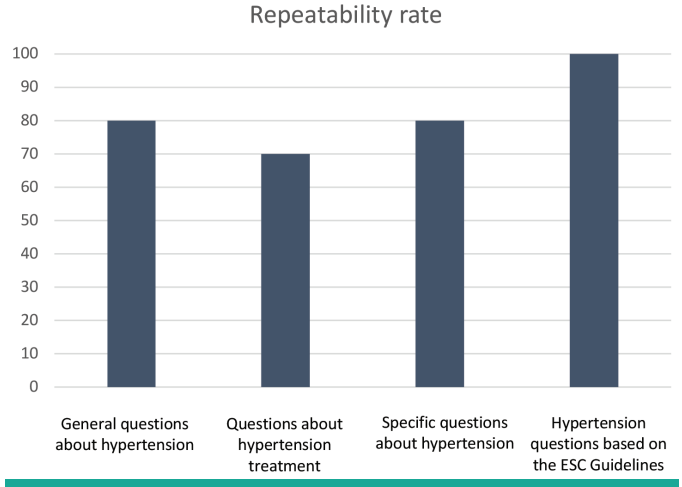


Figure 2. Repeatability rates of question answers
ESC: European Society of Cardiology

HT and/or some cardiological diseases have been addressed in different studies in the literature in terms of the functionality of AI (14-16). In an evaluation of the 100 HT questions, Almagazzachi et al. (17) reported that ChatGPT responses were appropriate in 92.5% of cases and inappropriate in 7.5%, with a reproducibility rate of 93% for ChatGPT. Kerkütlüoğlu et al. (18) evaluated ChatGPT in terms of knowledge and disease management regarding pulmonary HT, and 10 experts evaluated the performance of ChatGPT. Accordingly, the responses generated by ChatGPT were found to be reliable, with an average score of 8.4 (7.7-9.2), and valuable, with an average score of 7.9 (7.4-8.2).

In a study of laboratory and demographic data of 40 HT patients treated in a rural clinic in Georgia, Al Tibi et al. (19) compared the medical recommendations made by the cardiologist to the patients with the recommendations made by ChatGPT for the same patients and laboratory data. In contrast to the other studies mentioned above, discrepancies between the cardiologist and GPT-4 regarding general recommendations occurred in 95% of the 40 patients, while only 10.2% of specific medication recommendations were consistent between the cardiologist and GPT-4. Furthermore, the cardiologist and GPT-4 did not agree on medication changes. The authors highlighted the existence of different optimal laboratory value ranges among patients, the lack of a holistic analysis of GPT-4, and the need to provide complementary information to the model as the reasons for this discrepancy.

Our study also evaluated specific questions about HT, the ESC guidelines. Accordingly, 4 (80%) of the answers provided the highest quality and comprehensiveness. The answers to the ESC-based questions were also 100% reproducible. In their study conducted in 2023, Kusunose et al. (20) examined the ability of ChatGPT to accurately answer clinical questions about the Japanese Society of Hypertension's Hypertension Management Guidelines (JSH 2019). Similar to our study, ChatGPT had an 80% accuracy rate in responses to clinical questions. ChatGPT's performance in terms of HT responses was similar across two different guidelines. That ChatGPT complied with the guidelines shows that clinicians can use this model for information in their daily practice.

We emphasize that our work differs from prior studies by evaluating ChatGPT-4o for using a patient-centered question set written entirely in Turkish, addressing a non-English language context that remains underrepresented in the current literature; including guideline-based questions derived from the 2024 European Society of Cardiology Hypertension Guidelines, allowing for a structured assessment of guideline concordance; systematically assessing reproducibility by repeating all questions on different days under controlled conditions.

Study Limitations

Our study has some limitations. First, our patient-centered approach in determining the questions made those questions inherently subjective, and a relatively small number of questions and reviewers were included. Second, there may be differences in approach among the medical professionals and cardiologists who served as raters assessing the accuracy and consistency of the answers and information provided by ChatGPT. Furthermore, questions submitted to ChatGPT were limited to those from the healthcare forums and hospital websites examined for HT. It remains unclear how many questions are optimal for evaluating ChatGPT. Asking questions only in Turkish may have introduced a language bias. ChatGPT is rapidly evolving, and reproducibility may change in future versions.

Conclusion

We demonstrated that ChatGPT produces significantly accurate and reproducible answers to a variety of medical questions related to HT. ChatGPT may provide reliable information about HT, but it is important to seek professional medical advice before making any decisions about HT. Despite the limitations of the study, ChatGPT may serve as a useful source of information for both patients and healthcare professionals by using carefully. ChatGPT can provide only general information and support; a healthcare professional can make specific recommendations based on a holistic assessment of the patient's individual characteristics and diagnostics. Studies including a larger number of questions and medical professionals will shed more light on the use of ChatGPT as a source of information about HT.

Ethics

Ethics Committee Approval: Ethical approval for the study was obtained from the Institutional Ethics Committee of Medicana International İstanbul Hospital (decision number: 2025/2031; date: 24.09.2025).

Informed Consent: Since no patient data were used, informed consent was not required and therefore was not obtained.

Footnotes

Authorship Contributions

Surgical and Medical Practices: B.B.B., Concept: E.B.K., Design: E.B.K., G.G., Data Collection or Processing: E.B.K., G.G., Analysis or Interpretation: E.B.K., B.B.B., Literature Search: B.B.B., G.G., Writing: E.B.K.

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

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

References

1. OpenAI. ChatGPT: optimizing language models for dialogue, 2023. Available from:https://openai.com/blog/chatgpt/.
2. Pugliese N, Wai-Sun Wong V, Schattenberg JM, Romero-Gomez M, Sebastiani G; NAFLD Expert Chatbot Working Group; et al. Accuracy, reliability, and comprehensibility of ChatGPT-generated medical responses for patients with nonalcoholic fatty liver disease. Clin Gastroenterol Hepatol. 2024;22(4):886-889.e5.
3. Goodman RS, Patrinely JR, Stone CA Jr, Zimmerman E, Donald RR, Chang SS, et al. Accuracy and reliability of chatbot responses to physician questions. JAMA Netw Open. 2023;6:e2336483.
4. Horiuchi D, Tatekawa H, Shimono T, Walston SL, Takita H, Matsushita S, et al. Accuracy of ChatGPT generated diagnosis from patient's medical history and imaging findings in neuroradiology cases. Neuroradiology. 2024;66(1):73-79.
5. Alabdulmohsen DM, Almahmudi MA, Alhashim JN, Almahdi MH, Alkishy EF, Almossabeh MJ, et al. Is ChatGPT a reliable source of patient information on asthma? Cureus. 2024;16(7):e64114.
6. Kuşcu O, Pamuk AE, Sütay Süslü N, Hosal S. Is ChatGPT accurate and reliable in answering questions regarding head and neck cancer? Front Oncol. 2023;13:1256459.
7. Beilby K, Hammarberg K. ChatGPT: a reliable fertility decision-making tool? Hum Reprod. 2024;39(3):443-447.
8. Tunçer G, Güçlü KG. How reliable is ChatGPT as a novel consultant in infectious diseases and clinical microbiology? Infect Dis Clin Microbiol. 2024;6(1):55-59.
9. Gaziano TA, Bitton A, Anand S, Weinstein MC; International Society of Hypertension. The global cost of nonoptimal blood pressure. J Hypertens. 2009;27(7):1472-1477.
10. GBD 2015 Risk Factors Collaborators. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990-2015: a systematic analysis for the global burden of disease study 2015. Lancet. 2016;388(10053):1659-1724. Erratum in: Lancet. 2017;389(10064):e1.
11. Hypertension. GoogleTrends. [Internet]. Available from:https://trends.google.com/trends/explore?date=today%205-y&q=%2Fm%2F0k95h&hl=tr.
12. McEvoy JW, McCarthy CP, Bruno RM, Brouwers S, Canavan MD, Ceconi C, et al. 2024 ESC Guidelines for the management of elevated blood pressure and hypertension. Eur Heart J. 2024;45(3):3912-4018. Erratum in: Eur Heart J. 2025;46(14):1300. Erratum in: Eur Heart J. 2025;46(45):4949
13. Layton AT. AI, machine learning, and ChatGPT in hypertension. Hypertension. 2024;81(4):709-716.
14. Andreadis K, Rodriguez DV, Zakreuskaya A, Chen J, Gonzalez J, Mann D. Bridging gaps with generative AI: enhancing hypertension monitoring through patient and provider insights. Stud Health Technol Inform. 2024;316:939-943.
15. Saeed A, AlShafea A, A F, Bin Saeed A. Pacemaker malfunction in a patient with congestive heart failure and hypertension. Cureus. 2023;15(2):e34574.
16. Alam SF, Thongprayoon C, Miao J, Pham JH, Sheikh MS, Garcia Valencia OA, et al. Advancing personalized medicine in digital health: the role of artificial intelligence in enhancing clinical interpretation of 24-h ambulatory blood pressure monitoring. Digit Health. 2025;11:20552076251326014.
17. Almagazzachi A, Mustafa A, Eighaei Sedeh A, Vazquez Gonzalez AE, Polianovskaia A, Abood M, et al. Generative artificial intelligence in patient education: ChatGPT takes on hypertension questions. Cureus. 2024;16(2):e53441.
18. Kerkütlüoğlu M, Kaya E, Gökmen R. Trustworthiness, value, danger, and readability of ChatGPT-generated responses to health questions related to pulmonary arterial hypertension. Cureus. 2024;16(10):e71472.
19. Al Tibi G, Alexander M, Miller S, Chronos N. A retrospective comparison of medication recommendations between a cardiologist and ChatGPT-4 for hypertension patients in a rural clinic. Cureus. 2024;16(3):e55789.
20. Kusunose K, Kashima S, Sata M. Evaluation of the accuracy of ChatGPT in answering clinical questions on the Japanese Society of Hypertension Guidelines. Circ J. 2023;87(7):1030-1033.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):246-254
DOI: 10.4274/BMB.galenos.2026.56338

Association Between Fatty Liver Index and Epicardial Fat Thickness in Young Adults

Genç Yetişkinlerde Yağlı Karaciğer İndeksi ile Epikardiyal Yağ Kalınlığı Arasındaki İlişki

Güney Sarıoğlu

Battalgazi State Hospital, Clinic of Cardiology, Malatya, Turkey

Abstract

Objective: Hepatic steatosis refers to the fatty degeneration of liver tissue, which is associated with an increased risk of cardiovascular morbidity and mortality. Patients diagnosed with hepatic steatosis have been shown to have greater epicardial adipose tissue thickness. An increase in epicardial fat thickness (EFT) correlates with heightened cardiovascular risks. The fatty liver index (FLI) serves as a non-invasive metric for evaluating hepatic steatosis. In our study, we sought to assess whether EFT could be predicted within a apparently healthy young adult population utilising FLI. Furthermore, if such a prediction proves feasible, it aims to facilitate early diagnosis of at-risk individuals and the implementation of preventive measures to decrease future cardiovascular morbidity and mortality by furnishing pertinent information to clinicians.

Method: We conducted an observational cross-sectional study involving 258 participants randomly selected from apparently healthy young adults aged 18 to 41. The cohort was divided into two groups based on EFT (<4 mm and ≥4 mm). We assessed participants' EFT values using echocardiographic examination. We measured the FLI utilising variables such as body mass index, waist circumference, serum gamma-glutamyl transferase, and triglyceride level.

Results: A multiple linear regression model was constructed utilising both stepwise and enter methods, incorporating variables that demonstrated significant correlation with EFT as well as the gender variable, which has been associated with EFT in previous studies. The model significantly predicted EFT [R=0.492, R^2=0.242, adjusted R^2=0.233, F (3, 250) =26.662, p<0.001], accounting for 23.3% of the variance. The standard error of the estimate was 11.041, and the Durbin-Watson statistic was 2.096. Among the predictors, FLI, age,

Öz

Amaç: Karaciğer steatozu, kardiyovasküler morbidite ve mortalite riskinin artmasıyla ilişkili olan karaciğer dokusunun yağlı dejenerasyonunu ifade eder. Karaciğer steatozu tanısı konulan hastalarda epikardiyal yağ dokusu kalınlığının daha fazla olduğu gösterilmiştir. Epikardiyal yağ kalınlığındaki (EFT) artış, artan kardiyovasküler risklerle ilişkilidir. Yağlı karaciğer indeksi (FLI), hepatik steatozu değerlendirmek için kullanılan non-invaziv bir ölçüt olarak hizmet eder. Çalışmamızda, FLI'yı kullanarak görünüşte sağlıklı genç yetişkin popülasyonunda EFT'nin öngörülebilir olup olmadığını değerlendirmeyi amaçladık. Ayrıca, böyle bir öngörünün mümkün olduğu kanıtlanırsa, klinisyenlere ilgili bilgileri sunarak risk altındaki bireylerin erken teşhisini ve gelecekteki kardiyovasküler morbidite ve mortaliteyi azaltmak için önleyici tedbirlerin uygulanmasını kolaylaştırmayı hedeflemektedir.

Yöntem: On sekiz ila 41 yaşları arasında görünüşte sağlıklı genç yetişkinlerden rastgele seçilen 258 katılımcıyı içeren gözlemsel bir kesitsel çalışma yürüttük. Kohort, EFT'ye göre (<4 mm ve ≥4 mm) iki gruba ayrıldı. Ekokardiyografik muayene kullanarak katılımcıların EFT değerlerini değerlendirdik. Vücut kitle indeksi, bel çevresi, serum gama-glutamil transferaz ve trigliserid düzeyi gibi değişkenleri kullanarak FLI'yı ölçtük.

Bulgular: Hem stepwise hem de enter yöntemleri kullanılarak, EFT ile anlamlı korelasyon gösteren değişkenlerin yanı sıra önceki çalışmalarda EFT ile ilişkilendirilen cinsiyet değişkenini de içeren bir çoklu doğrusal regresyon modeli oluşturuldu. Model, EFT'yi önemli ölçüde tahmin etti [R=0,492, R^2=0,242, düzeltilmiş R^2=0,233, F (3, 250) =26,662, p<0,001] ve varyansın %23,3'ünü açıkladı. Tahminin standart hatası 11,041 ve Durbin-Watson istatistiği 2,096 idi.

Address for Correspondence: Güney Sarıoğlu, MD, Battalgazi State Hospital, Clinic of Cardiology, Malatya, Turkey
E-mail: guneysarioglu@outlook.com **ORCID:** orcid.org/0000-0002-1049-1873
Received: 29.09.2025 **Accepted:** 06.06.2026 **Epub:** 10.06.2026 **Publication Date:** 24.06.2026

Cite this article as: Sarıoğlu G. Association between fatty liver index and epicardial fat thickness in young adults. Bagcilar Med Bull. 2026;11(2):246-254

Abstract

and platelet count were all significant independent predictors of EFT.

Conclusion: As a result, we found that EFT increased with FLI. We assert that forthcoming large-scale multicentre research involving healthy young adults will confirm that the FLI can help healthcare professionals identify at-risk individuals early and implement measures to reduce future cardiovascular morbidity and mortality.

Keywords: Cardiometabolic risk, epicardial fat thickness, fatty liver index, young adults

Öz

Yordayıcılar arasında FLI, yaş ve trombosit sayısı, EFT'nin tümüyle anlamlı bağımsız yordayıcılarıydı.

Sonuç: Sonuç olarak, EFT'nin FLI ile birlikte arttığını bulduk. Sağlıklı genç yetişkinleri içeren gelecekteki büyük ölçekli çok merkezli araştırmaların, FLI'nın sağlık profesyonellerinin risk altındaki bireyleri erken teşhis etmelerine ve gelecekteki kardiyovasküler morbidite ve mortaliteyi azaltmak için önlemler almalarına yardımcı olabileceğini doğrulayacağını iddia ediyoruz.

Anahtar kelimeler: Epikardiyal yağ kalınlığı, genç yetişkinler, yağlı karaciğer indeksi, kardiyometabolik risk

Introduction

Hepatic steatosis is characterized by the accumulation of triglycerides in hepatocytes above normal levels. It is generally defined as fat accounting for at least 5% of liver weight (1).

The fatty liver index (FLI) is a widely used, simple, and non-invasive score developed to assess the risk of hepatic steatosis. The FLI is calculated by an algorithm based on body mass index (BMI), waist circumference, triglyceride level, and gamma-glutamyl transferase values (2-4). High FLI is associated with cardiometabolic risk factors, including insulin resistance, abdominal and cardiac adiposity, hypertension, and hyperlipidemia, and is also independently associated with all-cause mortality and cardiovascular disease risk (5-8).

Epicardial fat is a visceral adipose tissue surrounding the heart's outer surface and lying between the myocardium and the pericardium. It is metabolically active and secretes proinflammatory substances that can directly affect the heart and blood vessels (9). The functional complexity of epicardial fat thickness (EFT) is unclear. However, it has been shown to affect cardiac morphology and function, preserving heart contractility and repolarization under physiological conditions (10). Increased EFT is associated with greater visceral adiposity and a higher risk of metabolic syndrome.

EFT can be easily measured using transthoracic echocardiography (11,12). EFT is also associated with the severity of liver steatosis and fibrosis (13). In various studies, EFT above 4 mm is considered thick in healthy individuals (14,15).

Studies have indicated that as EFT increases, the risk of non-alcoholic liver disease also rises, and there is a significant and independent association between FLI and EFT within the general population (16,17). There are no studies in the current literature focusing on healthy pediatric or young adult populations. Our study is the first to examine this age group. According to our hypothesis, EFT will increase with increasing FLI in the apparently healthy young adult population.

In our study, we aim to determine whether FLI, a non-invasive technique, can predict EFT, recognized as an independent predictor of cardiovascular morbidity and mortality, in an apparently healthy young adult population.

Should our findings indicate that such a prediction is feasible, we aim to furnish clinicians with pertinent information, thereby facilitating early identification of at-risk individuals and the implementation of measures to mitigate future cardiovascular morbidity and mortality.

Materials and Methods

Study Population

We conducted an observational, cross-sectional study that included 258 apparently healthy young adults between the ages of 18 and 41 who presented to Battalgazi State Hospital's Cardiology Clinic with complaints such as chest pain, shortness of breath, or palpitations, and who were considered healthy after physical examination, laboratory tests, and clinical evaluation.

A total of 258 participants were initially recruited for the study. However, 3 participants were excluded from the final analysis due to missing EFT scores. Consequently, statistical evaluations were conducted with the remaining 255 valid cases.

The cohort was divided into two groups based on EFT (<4 mm and ≥4 mm).

In our study, we included patients aged 18-41 years who had no history of chronic disease, were not using medications, and had no active infection, obesity, or cachexia.

We excluded patients who were taking active medication, who had a history of acute and/or chronic disease, were obese, or who were over 41 years of age. We excluded participants with laboratory measurements suggesting acute infection.

The research was conducted in accordance with the Declaration of Helsinki and was approved by the Local Ethics Committee of İnönü University, as documented in decision number 2025/8194 (protocol no: 2025/8194; date: 21-08-2025). Written and signed consent was obtained from each participant.

Laboratory Measurements

We assessed participants' EFT values by echocardiography using a 3.2-MHz transducer on a Philips Affinity 50 ultrasound system (Philips, Andover, MA, USA). We measured EFT in the two-dimensional parasternal long-axis view by positioning the M-mode cursor perpendicular to the aortic annulus, along the free wall of the right ventricle, at end-diastole. We measured the FLI using BMI, waist circumference, serum gamma-glutamyl transferase, and triglyceride level (2). We calculated the BMI as weight in kilograms divided by the square of the height in meters, expressed as kg/m^2. Using a flexible, non-metallic tape measure, we measured waist circumference at the level of the navel and immediately above the iliac crest after a normal exhalation. After an 8 to 10-hour fast, we collected blood samples from the right or left antecubital vein for analysis. We obtained the history of coronavirus disease-2019 (COVID-19) infection from medical records.

Statistical Analysis

Data were analysed using SPSS software (IBM Corp., IBM SPSS Statistics for Mac, version 27.0; Armonk, NY: IBM Corp., 2020). By analysing the data in their recorded state, we have avoided any potential errors that could have been introduced by imputing missing values. A descriptive analysis was conducted to characterize the study population. The normality of the variables was tested using the Kolmogorov-Smirnov test. For pairwise comparisons, the independent-samples t-test was used for normally distributed variables and the Mann-Whitney U test for non-normally distributed variables. In this study, we considered results to be statistically significant at $p<0.05$.

Furthermore, 95% confidence intervals for differences that did not include zero were deemed statistically significant. The categorical variables were analysed using the appropriate chi-square test and were expressed as percentages and absolute numbers. Means and standard deviations were used to express continuous variables for normally distributed data, while medians and interquartile ranges were used for non-normally distributed data. The correlation between continuous variables was evaluated using Pearson's correlation test for normally distributed variables and Spearman's correlation test for non-normally distributed variables. Additionally, multiple linear regression analysis was conducted to forecast the EFT. We assessed the study's effect size using Cohen's d. A receiver operating characteristic (ROC) analysis evaluated the FLI's efficacy in predicting significant EFT.

Results

The study included 258 participants with a mean age of 29.31±6.86 years, of whom 39.50% were male (Table 1). When participants were stratified by EFT (EFT <4 mm vs. EFT ≥4 mm), the higher EFT group demonstrated significantly greater age (32.07±6.39 vs. 27.31±6.51 years, $p<0.001$), BMI (26.64±4.25 vs. 23.69±4.18 kg/m^2, $p<0.001$), and FLI (43.92±29.83 vs. 22.59±23.33, $p<0.001$) (Table 2).

Laboratory analysis showed significantly higher C-reactive protein (CRP) levels [0.20 (0.10-0.40) vs. 0.20 (0.10-0.30) mg/dL, $p=0.009$] (although median values were similar, the distribution differed significantly), triglycerides [102 (72-151) vs. 81.50 (63.50-116.50) mg/dL, $p=0.001$], and low-density lipoprotein (LDL) (112.97±28.28 vs. 99.63±28.88 mg/dL, $p<0.001$) in the higher EFT group.

No statistically significant differences were found between the groups regarding history of COVID-19 infection, haemoglobin, platelet count, total leukocyte count, monocyte count, basophil count, platelet distribution width, red cell distribution width, or thyroid-stimulating hormone levels ($p>0.05$ for all).

Statistically significant positive correlations between EFT and the variables were found using Pearson or Spearman correlation analysis. The strongest positive correlations were observed with age ($r=0.376$, $p<0.001$) and FLI ($r=0.413$, $p<0.001$) (Figure 1). Additionally, statistically significant correlations were identified between EFT and LDL ($r=0.255$, $p<0.001$), platelet count ($r=0.140$, $p=0.026$), CRP ($\rho=0.180$, $p=0.004$), and eosinophils ($\rho=0.136$, $p=0.030$). A weak

positive statistical relationship was observed between the EFT and the FLI (Figure 2).

We constructed a multiple linear regression model using stepwise and enter methods, including variables that showed significant correlations with EFT, together with the gender variable, which has been associated with EFT in previous studies. The model significantly predicted EFT [R=0.492, R^2=0.242, adjusted R^2=0.233, F (3, 250)=26.662, p<0.001], accounting for 23.3% of the variance. The standard error of the estimate was 11.041, and the Durbin-Watson statistic was 2.096. Among the predictors, age, FLI, and platelet count were all significant independent predictors of EFT (Table 3).

Cohen's d was used to assess the magnitude of differences between groups. The study revealed moderate-to-large effect sizes for age, BMI, and FLI, with d =0.74, 0.70, and 0.81, respectively.

Table 1. Demographic and laboratory data of participants

Characteristics (258 participants)	n (%)/mean (SD) or median (IQR)
Age (years) (SD)	29.31 (6.86)
Male n (%)	102 (39.51)
Smoking n (%)	98 (38)
COVID-19 (%)	92 (35.71)
BMI kg/m² (SD)	24.91 (4.44)
Fatty liver index (SD)	31.75 (28.38)
Haemoglobin g/dL (SD)	13.89 (1.81)
Platelet 10^3/mL (SD)	276.78 (61.74)
White blood cell 10^3/mL (SD)	7.31 (1.69)
Eosinophil 10^3/mL (IQR)	0.13 (0.08-0.22)
Monocyte 10^3/mL (SD)	0.57 (0.18)
Basophil 10^3/mL (IQR)	0.03 (0.02-0.04)
PDW % (SD)	12.01 (2.04)
RDW % (IQR)	12.75 (12.20-13.40)
TSH mU/L (IQR)	1.68 (1.24-2.21)
CRP mg/dL (IQR)	0.20 (0.10-0.40)
Triglyceride mg/dL (IQR)	90 (67-133.75)
LDL mg/dL (SD)	105.21 (29.58)
Epicardial fat thickness mm (SD)	4.15 (1.26)

The categorical variables were expressed as percentages (%) and absolute numbers. Means and standard deviations were used to express continuous variables for normally distributed data, while medians and interquartile ranges were used for non-normally distributed data. BMI: Body mass index, PDW: Platelet distribution width, RDW: Red-cell distribution width, TSH: Thyroid-stimulating hormone, CRP: C-reactive protein, LDL: Low-density lipoprotein, SD: Standard deviation, COVID-19: Coronavirus disease-2019, GFR: Glomerular filtration rate, IQR: Interquartile range. The demographic data is based on 258 individuals, whilst the epicardial fat thickness analyses are based on 255 individuals due to data loss

ROC analysis assessed the FLI's ability to predict substantial EFT. The results demonstrated a statistically significant area under the curve [(AUC) =0.727; standard error =0.032; 95% confidence interval (0.665, 0.789); p<0.001], indicating moderate discriminative ability (Figure 3). The analysis revealed that an FLI threshold of 14 yielded the highest Youden index (0.498), providing moderate discriminative performance, with sensitivity of 79.9% and specificity of 69.9% (18).

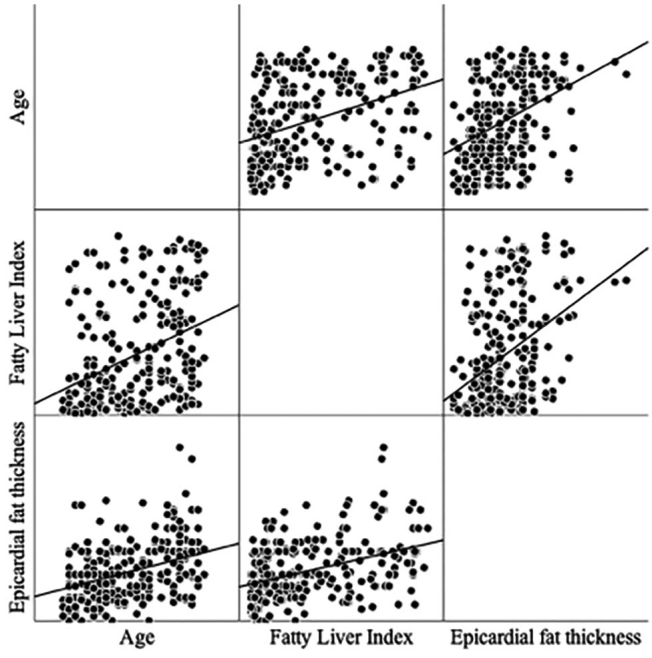


Figure 1. Scatterplot matrix showing the relationships between epicardial fat thickness, age, and fatty liver index

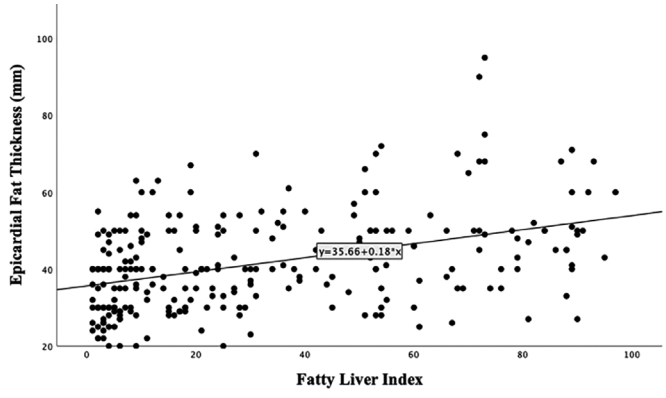


Figure 2. Correlation between epicardial fat thickness and fatty liver index

Table 2. Comparison of demographic and clinical data of both groups

Characteristics	EFT <4 mm (n=146)	EFT ≥4 mm (n=109)	p
Age (years) (SD)	27.31 (6.51)	32.07 (6.39)	<0.001**
Male n (%)	50 (34.25)	49 (44.95)	0.083
Smoking n (%)	34 (23.29)	62 (56.88)	0.069
COVID-19 n (%)	49 (33.56)	42 (38.53)	0.412
BMI kg/m^2 (SD)	23.69 (4.18)	26.64 (4.25)	<0.001**
Fatty liver index (SD)	22.59 (23.33)	43.92 (29.83)	<0.001**
Haemoglobin g/dL (SD)	13.73 (1.81)	14.04 (1.77)	0.127
Platelet 10^3/mL (SD)	271 (55.11)	284.31 (70.21)	0.102
White blood cell 10^3/mL (SD)	7.27 (1.71)	7.41 (1.71)	0.543
Eosinophil 10^3/mL (IQR)	0.13 (0.08-0.21)	0.14 (0.09-0.22)	0.140
Monocyte 10^3/mL (SD)	0.56 (0.18)	0.59 (0.19)	0.148
Basophil 10^3/mL (IQR)	0.03 (0.02-0.04)	0.03 (0.02-0.04)	0.227
PDW % (IQR)	12.05 (10.75-13.60)	11.70 (10.35-13.20)	0.078
RDW % (IQR)	12.75 (12.20-13.50)	12.80 (12.20-13.25)	0.604
TSH mU/L (IQR)	1.70 (1.21-2.34)	1.57 (1.21-2.12)	0.549
CRP mg/dL (IQR)	0.20 (0.10-0.30)	0.20 (0.10-0.40)	0.009*
Triglyceride mg/dL (IQR)	81.50 (63.50-116.50)	102 (72-151)	0.001**
LDL mg/dL (SD)	99.63 (28.88)	112.97 (28.28)	<0.001**

The categorical variables were analysed using the appropriate chi-square test and expressed as percentages (%) and absolute numbers. Means and standard deviations were used to express continuous variables for normally distributed data, while medians and interquartile ranges were used for non-normally distributed data. BMI: Body mass index, PDW: Platelet distribution width, RDW: Red-cell distribution width, TSH: Thyroid-stimulating hormone, SD: Standard deviation, COVID-19: Coronavirus disease-2019, CRP: C-reactive protein, LDL: Low-density lipoprotein, IQR: Interquartile range, *: p<0.05, **: p<0.001. The demographic data is based on 258 individuals, whilst the EFT analyses are based on 255 individuals due to data loss

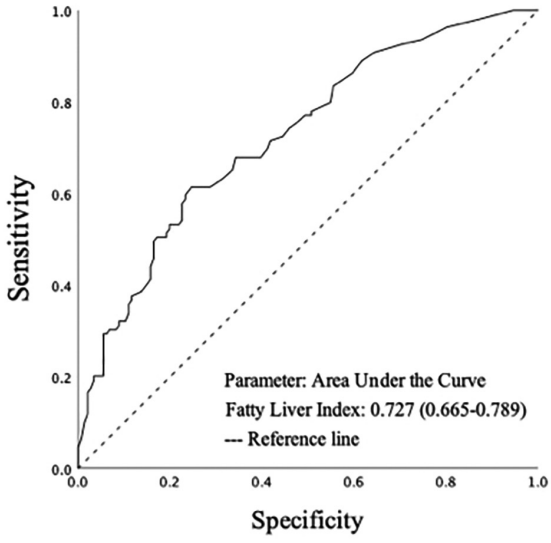


Figure 3. Receiver operating characteristic curve showing the fatty liver index's ability to discriminate between individuals with and without high epicardial fat thickness

Discussion

Our study is the first to examine the relationship between FLI and EFT in young, apparently healthy adults. We determined that as the FLI increases, the EFT also increases. Our multiple regression analysis reveals that the FLI is a significant independent predictor of EFT.

Hepatic steatosis is characterized by excessive fat accumulation, particularly triglycerides, in hepatocytes (liver cells) and is often called fatty liver disease. Hepatosteatosis is the first and most common stage of both alcohol-related and non-alcohol-related liver diseases (19,20). Hepatic steatosis often co-occurs with cardiovascular disease. It is associated with higher coronary artery calcium scores, greater plaque burden, and metabolic syndrome in patients (21,22). Hepatosteatosis is associated with cardiovascular disease in the general population (23). FLI is a score developed to estimate hepatic steatosis non-invasively, based on BMI, waist circumference, triglycerides, and gamma-glutamyltransferase levels. FLI is widely used to screen for the risk of non-alcoholic fatty liver disease and to predict the presence of steatosis in population studies (24).

Table 3. Multiple linear regression analysis predicting epicardial fat thickness using the enter approach

	B	SE	β	t	p	95% CI	
						Lower	Upper
Age	0.48	0.11	0.26	4.37	<0.01	0.26	0.69
Fatty liver index	0.14	0.03	0.3	5.09	<0.01	0.08	0.19
Platelet count	0.03	0.01	0.13	2.4	0.02	0.01	0.05

The dependent variable is the epicardial fat tissue thickness. CI: Confidence interval, SE: Standard error

Epicardial adipose tissue (EAT) is visceral adipose tissue located around the heart and coronary arteries, between the myocardium and the pericardium. It shares the same microcirculation as the heart and is in close anatomical proximity to it. Under normal conditions, EAT performs protective functions such as cushioning the heart against mechanical stress, regulating myocardial temperature, and maintaining fatty acid balance. It also serves as an energy source and secretes various cytokines that nourish the heart muscle (25).

EFT is typically measured by echocardiography or CT and serves as an independent indicator of visceral obesity, metabolic syndrome, type 2 diabetes, hypertension, and increased cardiovascular disease risk (26). Reviews and meta-analyses show that EFT and cardiovascular risk increase in conditions such as fatty liver disease and metabolic syndrome (27,28).

Research suggests that an increase in hepatic steatosis is associated with a corresponding increase in EFT, thereby increasing the risk of cardiometabolic complications. This connection is evident in conditions such as diabetes, obesity, and metabolic syndrome (29,30).

The FLI is a standard biochemical score that indicates the level of liver steatosis. Research shows that as FLI rises, EFT also increases significantly. Findings from an extensive study involving 512 patients with non-alcoholic fatty liver disease demonstrated a significant and independent association between FLI and EFT. Multivariate analysis indicated that higher FLI was associated with greater EFT. Additionally, as the severity of liver steatosis escalated, EFT also showed a corresponding rise (17).

In recent years, studies have examined the relationship between EFT and carotid intima-media thickness. Karakurt et al. (31) found a significant relationship between the presence and severity of erectile dysfunction in newly diagnosed hypertensive patients and EFT and carotid intima-media thickness, and Ardahanlı et al. (32) found significant reductions in EFT and carotid intima-media thickness with empagliflozin treatment in patients with diabetes mellitus.

No studies have directly examined the relationship between the FLI and EFT in apparently healthy young adults. Our study is the first to examine this issue.

In our study, FLI, age, and platelet count were predictive of EFT. The possible physiological reasons for this are as follows:

a) FLI

Both fatty liver and epicardial fat accumulation are indicators of insulin resistance and metabolic syndrome. Insulin resistance causes fat to accumulate in both the liver and around the heart (16). Ectopic fat tissues (such as hepatic and epicardial fat) increase systemic and local inflammation by secreting proinflammatory cytokines and adipokines. This condition increases both liver damage and cardiovascular risk (33).

b) Age

With ageing, body fat distribution changes and visceral fat, especially epicardial fat, increases. This increase contributes to an elevated cardiometabolic risk (34). Ectopic accumulation of adipose tissue in organs such as muscles, liver, and heart increases with age. Fibrotic and apoptotic changes in epicardial fat also become more pronounced with ageing (35).

c) Platelet count

EAT contributes to heightened local and systemic inflammatory responses by secreting proinflammatory cytokines and adipokines. Elevated inflammation may lead to increased platelet activation and platelet count (36). In conditions such as metabolic syndrome, obesity, and insulin resistance, both platelet count and EFT are elevated. This common increase is indicative of underlying inflammatory and metabolic disorders (26).

In our study, the echocardiographic EFT cut-off value was set at 4 mm. Previous studies have reported a wide range of EFT thresholds depending on the population studied

and the clinical outcome assessed. In patients with acute ischaemic stroke, ROC analysis identified 3.75 mm as the optimal cut-off value, while a threshold of 4 mm was adopted for practical clinical use (37). In contrast, studies evaluating coronary artery disease severity have generally reported higher cut-off values, ranging from 4.5 to 5.5 mm, for predicting the presence of coronary artery disease or more extensive multivessel involvement (38). Similarly, an EFT value of ≥5 mm has been associated with increased coronary artery disease risk in diabetic populations (39). Therefore, the selection of a 4 mm cut-off in our study is consistent with thresholds reported in acute ischaemic stroke populations while remaining below those described in higher-risk cardiovascular cohorts, potentially providing a more sensitivity-oriented approach for risk stratification (37-39).

Study Limitations

Our study has some limitations. Due to its cross-sectional design, we can't establish a cause-and-effect relationship. Instead, the study shows only the relationship between the variables at a single point in time, without explaining how they interact. Although the participants were deemed healthy based on their physical examination, laboratory results, and clinical assessment, they were apparently healthy but symptomatic young adults, which may limit the generalisability to community-dwelling asymptomatic young adults. Limitations of generalizability: The study population was limited to young adults aged 18-41. Therefore, the findings cannot be applied to older populations or to children. This limits the study's generalisability. The explanatory power of the multivariate regression model used in our study was limited. This suggests that the factors determining EFT cannot be explained solely by the variables included in our study and that additional metabolic, lifestyle, and behavioural determinants should also be considered. Since FLI incorporates BMI and triglycerides, which are themselves related to EFT, the observed association between FLI and EFT may be partly driven by these shared components. Therefore, our findings should be interpreted to indicate a composite cardiometabolic risk profile captured by FLI, rather than a fully independent effect of liver steatosis on EFT. The literature indicates that EFT is closely associated with insulin resistance and impaired glucose metabolism, and that EFT is significantly increased in both obese individuals and non-diabetic insulin-resistant patients (40). Excluding parameters that directly reflect insulin resistance from our model may have contributed to unexplained variance in EFT. EFT is influenced not only by metabolic parameters but also by lifestyle factors. A cohort study conducted in the general population showed that physical activity is inversely related to EFT, while red meat consumption is directly related; alcohol consumption and heavy drinking have been reported to be associated with increased EFT, particularly in women (41). The absence of variables such as physical activity level, dietary pattern, and alcohol consumption in our analyses may significantly limit the model's explanatory power. Therefore, the limited explanatory power of our regression model indicates that additional variables, such as well-defined indicators of insulin resistance affecting EFT, detailed physical activity measurements, and dietary habits, should be integrated into future studies. The study indicated that the optimal threshold value for the FLI to predict significant EFT is 14. Although this value offers the best performance, characterized by a high Youden index (0.498) and a moderate AUC of 0.727, further research is needed to assess its effectiveness and reliability in clinical practice.

Conclusion

We found that EFT was positively associated with FLI. In addition, the participant group with higher EFT was significantly older and had a significantly higher BMI and FLI than the group with lower EFT. In this group, levels of CRP, triglycerides, and LDL were also significantly higher. Longitudinal follow-up studies may be conducted to determine whether FLI is associated with EFT in healthy young adults.

We assert that forthcoming extensive multicenter research involving healthy young adults will substantiate that the FLI can aid healthcare professionals in early identification of at-risk individuals and in the implementation of measures to reduce future cardiovascular morbidity and mortality.

Ethics

Ethics Committee Approval: The research was conducted in accordance with the Declaration of Helsinki and was approved by the Local Ethics Committee of İnönü University, as documented in decision number 2025/8194 (protocol no: 2025/8194; date: 21-08-2025).

Informed Consent: Written and signed consent was obtained from each participant.

Acknowledgments

I would like to thank the participants who contributed to this study.

Footnotes

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

References

1. Nassir F, Rector RS, Hammoud GM, Ibdah JA. Pathogenesis and prevention of hepatic steatosis. Gastroenterol Hepatol (N Y). 2015;11(3):167-175.
2. Bedogni G, Bellentani S, Miglioli L, Masutti F, Passalacqua M, Castiglione A, et al. The fatty liver index: a simple and accurate predictor of hepatic steatosis in the general population. BMC Gastroenterol. 2006;6:33.
3. Kaneva AM, Bojko ER. Fatty liver index (FLI): more than a marker of hepatic steatosis. J Physiol Biochem. 2024;80(1):11-26.
4. Koehler EM, Schouten JN, Hansen BE, Hofman A, Stricker BH, Janssen HL. External validation of the fatty liver index for identifying nonalcoholic fatty liver disease in a population-based study. Clin Gastroenterol Hepatol. 2013;11(9):1201-1204.
5. Carli F, Sabatini S, Gaggini M, Sironi AM, Bedogni G, Gastaldelli A. Fatty liver index (FLI) identifies not only individuals with liver steatosis but also at high cardiometabolic Risk. Int J Mol Sci. 2023;24(19):14651.
6. Kim JH, Moon JS, Byun SJ, Lee JH, Kang DR, Sung KC, et al. Fatty liver index and development of cardiovascular disease in Koreans without pre-existing myocardial infarction and ischemic stroke: a large population-based study. Cardiovasc Diabetol. 2020;19(1):51.
7. Zou B, Yeo YH, Cheung R, Ingelsson E, Nguyen MH. Fatty liver index and development of cardiovascular disease: findings from the UK biobank. Dig Dis Sci. 2021;66(6):2092-2100.
8. Chung GE, Jeong SM, Cho EJ, Yoo JJ, Cho Y, Lee KN, et al. Association of fatty liver index with all-cause and disease-specific mortality: a nationwide cohort study. Metabolism. 2022;133:155222.
9. Salazar J, Luzardo E, Mejías JC, Rojas J, Ferreira A, Rivas-Ríos JR, et al. Epicardial fat: physiological, pathological, and therapeutic implications. Cardiol Res Pract. 2016;2016:1291537.
10. Iacobellis G, Leonetti F, Singh N, M Sharma A. Relationship of epicardial adipose tissue with atrial dimensions and diastolic function in morbidly obese subjects. Int J Cardiol. 2007;115(2):272-273.
11. Iacobellis G, Willens HJ, Barbaro G, Sharma AM. Threshold values of high-risk echocardiographic epicardial fat thickness. Obesity (Silver Spring). 2008;16(4):887-892.
12. Iacobellis G, Willens HJ. Echocardiographic epicardial fat: a review of research and clinical applications. J Am Soc Echocardiogr. 2009;22(12):1311-1319; quiz 1417-1418.
13. Petta S, Argano C, Colomba D, Cammà C, Di Marco V, Cabibi D, et al. Epicardial fat, cardiac geometry and cardiac function in patients with non-alcoholic fatty liver disease: association with the severity of liver disease. J Hepatol. 2015;62(4):928-933.
14. Swaroop S, Tripathy SK, Sahu NC, Das B, Satpathy M. Epicardial fat thickness and its association as a risk factor for coronary artery disease. World Journal of Cardiovascular Diseases. 2019;9(3):193-201.
15. Shambu SK, Desai N, Sundaresh N, Babu MS, Madhu B, Gona OJ. Study of correlation between epicardial fat thickness and severity of coronary artery disease. Indian Heart J. 2020;72(5):445-447.
16. Orci LA, Jornayvaz FR, Toso C, Gariani K. Systematic review and meta-analysis of the usefulness of epicardial fat thickness as a non-invasive marker of the presence and severity of nonalcoholic fatty liver disease. Biomedicines. 2022;10(9):2204.
17. Fracanzani AL, Pisano G, Consonni D, Tiraboschi S, Baragetti A, Bertelli C, et al. Epicardial adipose tissue (EAT) thickness is associated with cardiovascular and liver damage in nonalcoholic fatty liver disease. PLoS One. 2016;11(9):e0162473.
18. Fluss R, Faraggi D, Reiser B. Estimation of the Youden index and its associated cutoff point. Biom J. 2005;47(4):458-472.
19. An L, Wirth U, Koch D, Schirren M, Drefs M, Koliogiannis D, et al. The role of gut-derived lipopolysaccharides and the intestinal barrier in fatty liver diseases. J Gastrointest Surg. 2022;26(3):671-683.
20. Viscarra J, Sul HS. Epigenetic regulation of hepatic lipogenesis: role in hepatosteatosis and diabetes. Diabetes. 2020;69(4):525-531.
21. Carter J, Heseltine TD, Meah MN, Tzolos E, Kwiecinski J, Doris M, et al. Hepatosteatosis and atherosclerotic plaque at coronary CT angiography. Radiol Cardiothorac Imaging. 2022;4(2):e210260.
22. Choi KM, Han K, Park S, Chung HS, Kim NH, Yoo HJ, et al. Implication of liver enzymes on incident cardiovascular diseases and mortality: a nationwide population-based cohort study. Sci Rep. 2018;8(1):3764.
23. Heseltine T, Murray S, Ortega Martorell S, Olier I, Lip GYH, Khoo S. Associations of hepatosteatosis with cardiovascular disease in HIV-positive and HIV-negative patients: the Liverpool HIV-heart project. J Acquir Immune Defic Syndr. 2021;87(5):1221-1227.
24. Castellana M, Donghia R, Guerra V, Procino F, Lampignano L, Castellana F, et al. Performance of fatty liver index in identifying non-alcoholic fatty liver disease in population studies. A meta-analysis. J Clin Med. 2021;10(9):1877.
25. Iacobellis G. Epicardial adipose tissue in contemporary cardiology. Nat Rev Cardiol. 2022;19(9):593-606.
26. Chong B, Jayabaskaran J, Ruban J, Goh R, Chin YH, Kong G, et al. Epicardial adipose tissue assessed by computed tomography and echocardiography are associated with adverse cardiovascular outcomes: a systematic review and meta-analysis. Circ Cardiovasc Imaging. 2023;16(5):e015159.
27. Villasante Fricke AC, Iacobellis G. Epicardial adipose tissue: clinical biomarker of cardio-metabolic risk. Int J Mol Sci. 2019;20(23):5989.
28. Iacobellis G. Local and systemic effects of the multifaceted epicardial adipose tissue depot. Nat Rev Endocrinol. 2015;11(6):363-371.
29. Aitken-Buck HM, Moharram M, Babakr AA, Reijers R, Van Hout I, Fomison-Nurse IC, et al. Relationship between epicardial adipose tissue thickness and epicardial adipocyte size with increasing body mass index. Adipocyte. 2019;8(1):412-420.
30. Gruzdeva O, Uchasova E, Dyleva Y, Borodkina D, Akbasheva O, Belik E, et al. Relationships between epicardial adipose tissue thickness and adipo-fibrokine indicator profiles post-myocardial infarction. Cardiovasc Diabetol. 2018;17(1):40.
31. Karakurt D, Ede H, Turan Y, Gürel A, Ardahanlı İ, Sarak T. Yeni tanı konmuş hipertansiyonu olan hastalarda erektil disfonksiyon ile epikardiyal yağ dokusu kalınlığı ve karotis intima-media kalınlığı arasındaki ilişkinin değerlendirilmesi [Evaluation of the relationship between erectile dysfunction and epicardial fat tissue thickness and carotid intima-media thickness in patients with newly diagnosed hypertension]. Turk Kardiyol Dern Ars. 2020;48(8):746-753.

32. Ardahanlı İ, Aslan R, Çelik M, Akgün O, Akyüz O. Effects of empagliflozin on carotid intima-media thickness and epicardial fat tissue volume in patients with type-2 diabetes mellitus. Lokman Hekim Health Sci. 2021;1(3):74-80.

33. Petta S, Craxì A. Epicardial fat in patients with non-alcoholic fatty liver disease. J Hepatol. 2015;62(5):1215.

34. Chen YC, Lee WH, Lee MK, Hsu PC, Tsai WC, Chu CY, et al. Epicardial adipose tissue thickness is not associated with adverse cardiovascular events in patients undergoing haemodialysis. Sci Rep. 2020;10(1):6281.

35. Iacobellis G. Aging effects on epicardial adipose tissue. Front Aging. 2021;2:666260.

36. Ustuntas G, Basat SU, Calik AN, Sivritepe R, Basat O. Relationship between epicardial fat tissue thickness and CRP and neutrophil-lymphocyte ratio in metabolic syndrome patients over 65 years. Sisli Etfal Hastan Tip Bul. 2021;55(3):405-411.

37. Qayyum W, Iqbal M, Khan S, Yousafzai Z, Nabi D, Ashraf S, et al. Epicardial fat thickness as a predictor of acute ischemic stroke at Rehman medical institute: a case-control study. J Rehman Med Inst. 2025;11(1):19-23.

38. El-Naggar HM, Abdel-Maseh JG, Hasan-Ali H, Khidr SS. Epicardial fat thickness predicts severe coronary artery disease and high mortality risk among ST-elevation myocardial infarction patients. Echo Res Pract. 2025;12(1):17.

39. Akhter AN, Aisha F, Moazzam AB, Khan SHB, Malik J, Parveen A. Epicardial fat thickness as a marker of coronary artery disease in diabetics: a single center study. Clin Cardiol. 2025;48(7):e70171.

40. Altin C, Sade LE, Gezmis E, Yilmaz M, Ozen N, Muderrisoglu H. Assessment of epicardial adipose tissue and carotid/femoral intima media thickness in insulin resistance. J Cardiol. 2017;69(6):843-850.

41. Gustafsson B, Rovio SP, Ruohonen S, Hutri-Kähönen N, Kähönen M, Viikari JSA, et al. Determinants of echocardiographic epicardial adipose tissue in a general middle-aged population - The Cardiovascular Risk in Young Finns Study. Sci Rep. 2024;14(1):11982.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):255-260
DOI: 10.4274/BMB.galenos.2026.42714

A Retrospective Analysis of Colposcopy Outcomes in High-risk HPV-positive Individuals Referred to Our Clinic

Kliniğimize Yönlendirilen Yüksek Riskli HPV Pozitif Hastaların Kolposkopi Sonuçlarının Retrospektif Analizi

Erkan Şimşek[1], Gökşen Görgülü[1], Mustafa Gökkaya[1], Sadık Gündüz[2], İbrahim Taşkum[3]

[1]University of Health Sciences Turkey, Gaziantep City Hospital, Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Gaziantep, Turkey
[2]University of Health Sciences Turkey, Bakırköy Dr. Sadi Konuk Training and Research Hospital, Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, İstanbul, Turkey
[3]University of Health Sciences Turkey, Gaziantep City Hospital, Department of Obstetrics and Gynecology, Gaziantep, Turkey

Abstract

Objective: Cervical cancer accounted for 2.4% of all newly diagnosed malignancies among women in Turkey, with approximately 1.588 new cases and 722 deaths reported in 2020. Current guidelines recommend a risk-based management approach in which women with high-risk human papilloma virus (HPV) positivity, regardless of cytology results, are often referred for colposcopic evaluation. This study aimed to compare our colposcopic biopsy findings in high-risk HPV-positive patients with the existing literature.

Method: This retrospective study included individuals referred to our tertiary care center between February 2024 and May 2025 due to high-risk HPV positivity, in line with current clinical guidelines. Data on patients' age, menopausal status, smoking habits, cytology results, HPV genotypes, and colposcopic biopsy outcomes were recorded and analyzed.

Results: A total of 180 patients were included. Notably, no cases of adenocarcinoma or invasive cervical cancer were detected. There were no statistically significant associations between age, menopausal status, smoking status, or HPV genotype and colposcopic biopsy outcomes (normal, CIN1, CIN2, CIN3). The only statistically significant finding was that CIN3 lesions were more frequently observed in women aged 51-65 years compared to other age groups (p=0.034).

Öz

Amaç: Türkiye'de serviks kanseri, kadınlarda yeni tanı konulan tüm malignitelerin %2,4'ünü oluşturmaktadır ve 2020 yılında yaklaşık 1,588 yeni olgu ile 722 ölüm bildirilmiştir. Güncel kılavuzlar, sitoloji sonucundan bağımsız olarak yüksek riskli insan papilloma virüsü (HPV) pozitifliği saptanan kadınların kolposkopi ile değerlendirilmesini öneren risk temelli bir yönetim yaklaşımını benimsemektedir. Bu çalışma, yüksek riskli HPV pozitif bireylerde elde ettiğimiz kolposkopik biyopsi bulgularını mevcut literatürle karşılaştırmayı amaçlamaktadır.

Yöntem: Bu retrospektif çalışmada, Şubat 2024 ile Mayıs 2025 tarihleri arasında yüksek riskli HPV pozitifliği nedeniyle üçüncü basamak sağlık merkezimize yönlendirilen hastalar, güncel klinik kılavuzlar doğrultusunda değerlendirildi. Hastaların yaşı, menopozal durumu, sigara kullanımı, smear sonuçları, HPV tipleri ve kolposkopik biyopsi sonuçları kaydedilerek analiz edildi.

Bulgular: Çalışmaya toplam 180 hasta dahil edildi. Önemli olarak, adenokarsinom ya da invaziv servikal kanser saptanmadı. Yaş, menopoz durumu, sigara kullanımı ve HPV tipleri ile kolposkopik biyopsi sonuçları (normal, CIN1, CIN2, CIN3) arasında istatistiksel olarak anlamlı bir ilişki bulunmadı. Sadece 51-65 yaş grubundaki hastalarda CIN3 lezyonlarının diğer yaş gruplarına göre anlamlı düzeyde daha sık görüldüğü belirlendi (p=0,034).

Address for Correspondence: Erkan Şimşek, MD, University of Health Sciences Turkey, Gaziantep City Hospital, Department of Obstetrics and Gynecology, Division of Gynecologic Oncology, Gaziantep, Turkey
E-mail: dr.erkann@hotmail.com **ORCID:** orcid.org/0000-0002-6723-1773
Received: 02.10.2025 **Accepted:** 04.04.2026 **Epub:** 20.04.2026 **Publication Date:** 24.06.2026

Cite this article as: Şimşek E, Görgülü G, Gökkaya M, Gündüz S, Taşkum İ. A retrospective analysis of colposcopy outcomes in high-risk HPV-positive individuals referred to our clinic. Bagcilar Med Bull. 2026;11(2):255-260

Abstract

Conclusion: Our colposcopic biopsy outcomes in high-risk HPV-positive patients were found to be consistent with the current literature.

Keywords: Cervical dysplasia, colposcopic biopsy, high-risk HPV positivity

Öz

Sonuç: Yüksek riskli HPV pozitif hastalarda elde ettiğimiz kolposkopik biyopsi sonuçları, mevcut literatür ile uyumlu bulunmuştur.

Anahtar kelimeler: Kolposkopik biyopsi, servikal displazi, yüksek riskli HPV pozitifliği

Introduction

Cervical cancer, a highly preventable malignancy by screening, resulted in the deaths of 4.138 women in the United States, averaging 11 fatalities per day, with half of the deceased aged 58 years or younger (1). Marked inequalities in the incidence and mortality rates of cervical cancer were noted. The highest incidence rates are observed in Sub-Saharan Africa, with 85% of fatalities occurring in underdeveloped nations worldwide (2). GLOBOCAN 2020 projections indicate that cervical cancer ranks as the fourth most prevalent disease among women globally, with over 604,000 new cases and 342,000 fatalities recorded in 2020 (3). Nearly 90% of cervical cancer deaths occur in resource-limited settings, underscoring global disparities in cancer prevention and care. Cervical cancer constituted 2.4% of all newly diagnosed malignancies in women in Turkey, with about 1.588 new cases and 722 fatalities reported in 2020 (3). The age-standardized incidence rate was 4.3 per 100,000 women, quite low in comparison to worldwide norms, presumably attributable to national cervical cancer screening initiatives aimed at women aged 30-65 years (4). Nonetheless, sustained efforts are required to enhance participation rates and adherence to follow-up, particularly among high-risk demographics. Girls aged 9 to 14 are incorporated into the immunization program in Turkey.

Persistent infection with high-risk human papillomavirus (HR-HPV), particularly HPV 16 and HPV 18, accounts for approximately 70% of cervical cancer cases (5). The early detection of cervical intraepithelial neoplasia (CIN) and the reduction of invasive cervical cancer have been significantly improved in countries with robust healthcare systems as a result of the implementation of HPV DNA testing and cytology-based screening programs. Current guidelines emphasize, a risk-based management approach, where women with positive HPV results, regardless of cytology status, are often referred for colposcopic evaluation (6).

Recent studies indicate that even women with normal cytology but positive HPV 16/18 genotypes carry a substantial risk of underlying high-grade lesions (CIN2+), with detection rates ranging from 10% to 40% (7). Consequently, colposcopy-guided biopsy remains the gold standard for evaluating HPV-positive women and confirming the presence of precancerous or cancerous lesions.

The aim of this study was to evaluate colposcopy-guided biopsy outcomes in patients referred to our clinic with high-risk HPV positivity and to examine their association with cervical cytology findings.

Materials and Methods

This retrospective analysis encompassed patients who received colposcopic biopsy at University of Health Sciences Turkey, Gaziantep City Hospital from February 2024 to May 2025. The study included patients who had a colposcopic assessment after being found to be cervical HPV positive for the first time. The criteria for exclusion were as follows: Possessed a history of surgical intervention for CIN or cervical carcinoma, radiation, or chemotherapy, and had undergone colposcopy without cervical biopsy. The study was approved by the Medical Ethics Committee of University of Health Sciences Turkey, Gaziantep City Hospital (approval no: 221/2025, date: 21.05.2025).

The cervical cytology analysis was carried out using the liquid-based cytology test. Expert pathologists analyzed the data and categorized them using the 2014 Bethesda system (8). The Bethesda technique was utilized for cytological evaluation, producing the subsequent results: negative for intraepithelial lesions or malignancy (NILM—normal), atypical squamous cells of undetermined significance (ASC-US), atypical squamous cells—cannot exclude high-grade squamous intraepithelial lesion (ASC-H), low-grade intraepithelial lesion (LSIL), high-grade intraepithelial lesion (HSIL), atypical glandular cells (AGC), re-evaluation of insufficient cytology results was conducted. Kits from Qiagen were used to collect HPV DNA samples. The study included 14 carcinogenic HR-HPV types: HPV 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, 68, and 73 (9).

Colposcopic exams were performed in accordance with the 2020 American Society for Colposcopy and Cervical Pathology and Turkish Society of Gynecologic Oncology

Guidelines (10,11) and by gynecological oncologists. Cervical colposcopic examinations were performed using an EDAN C6A video colposcope (EDAN Instruments, Shenzhen, China) equipped with a green filter for enhanced visualization of vascular patterns. Multiple cervical biopsies were obtained from aberrant imaging locations. In the event of HR-HPV positive and the lack of concerning lesions in the colposcopic biopsy, a normal biopsy was conducted at positions 3, 6, 9, and 12. Endocervical curettage was performed in all patients who underwent colposcopic evaluation. The histopathology findings were categorized into four groups: Low-grade squamous intraepithelial lesions (comprising koilocytic changes, warts, and CIN I), high-grade squamous intraepithelial lesions [containing CIN II/III and carcinoma *in situ* (CIS)], chronic cervicitis or mucositis (inflammation), and no abnormality intraepithelial lesion or malignancy observed (normal).

Statistical Analysis

Statistical analyses were performed using SPSS software (version 25.0; IBM Corp., Armonk, NY, USA). Continuous variables were expressed as mean ± standard deviation, while categorical variables were presented as number and percentage [n (%)]. Comparisons between categorical variables were performed using the chi-square test. Univariate and multivariate logistic regression analyses were conducted to identify factors associated with CIN2+ lesions. Odds ratios and 95% confidence intervals were calculated. A p-value <0.05 was considered statistically significant.

Results

A total of 180 patients were referred to our clinic from family medicine units and other healthcare institutions (including private hospitals and secondary state hospitals) due to positive HPV test results. The mean age of the patients was 41.9±9.1 years. Among the study population, 80 (44.4%) patients were aged 30-40 years, 67 (37.2%) were aged 41-50 years, and 33 (18.5%) were aged 51-65 years. Overall, 44 (24.5%) patients were postmenopausal, and 56 (31.2%) were smokers. HPV16 was detected in 81 (45%) patients, HPV18 in 29 (16.1%) patients, and other high-risk HPV types in 40 (22.2%) patients. The distribution of other high-risk HPV types was as follows: HPV31 in 4 (2.2%), HPV33 in 6 (3.3%), HPV39 in 4 (2.2%), HPV45 in 4 (2.2%), HPV51 in 4 (2.2%), HPV52 in 4 (2.2%), HPV56 in 4 (2.2%), HPV58 in 4 (2.2%), and HPV68 in 6 (3.3%) patients.

Multiple HPV infection was detected in 30 (16.7%) patients. Dual HPV positivity included HPV16-18 in 8 (4.4%), HPV16-33 in 6 (3.3%), HPV31-33 in 5 (2.7%), HPV52-68 in 3 (1.6%), and HPV66-68 in 2 (1.1%) patients. Triple HPV positivity included HPV16-31-56 in 2 (1.1%), HPV39-45-53 in 1 (0.5%), HPV31-33-45 in 1 (0.5%), HPV31-33-66 in 1 (0.5%), and HPV52-66-68 in 1 (0.5%) patients. Cytology results showed that 127 (70.5%) patients had NILM, 32 (17.9%) had ASC-US, 9 (5%) had LSIL, 9 (5%) had HSIL, 2 (1.1%) had ASC-H, and 1 (0.5%) had AGC. Colposcopy-guided biopsy results revealed normal findings in 128 (71.1%) patients, CIN1 in 15 (8.3%), CIN2 in 9 (5%), and CIN3 in 28 (15.6%) patients. No cases of adenoCIS or invasive carcinoma were identified. The clinicopathologic characteristics of the patients are summarized in Table 1.

Table 1. Clinicopathologic characteristics of the patients

Patients characteristics	n=180	Percentage (%)
Age (mean ± SD), years old	41.9±9.1	
30-40 years	80	44.4
41-50 years	67	37.2
51-65 years	33	18.5
Menopausal status of patients		
Premenopause	136	75.5
Postmenopause	44	24.5
Patients' smoking status		
Non-smoker	124	68.8
Smoker	56	31.2
HR-HPV typing		
HPV16	81	45
HPV18	29	16.1
Non-16/18, HR-HPV	40	22.2
Multiple (dual) HPV Positivity	30	16.7
Cervical cytology		
NILM	127	70.5
ASC-US	32	17.9
LSIL	9	5
HSIL	9	5
ASC-H	2	1.1
AGC	1	0.5
Pathological biopsy result		
Normal	128	71.1
CIN 1	15	8.3
CIN 2	9	5
CIN 3	28	15.6

SD: Standard deviation, AGC: Abnormal glandular cells, AIS: Adenocarcinoma *in situ*, ASC-H: Atypical squamous cells—cannot exclude high-grade squamous intraepithelial lesion, ASC-US: Atypical squamous cells—undetermined significance, CIN: Cervical intraepithelial neoplasia, HR-HPV: High-risk human papillomavirus, HSIL: High-grade intraepithelial lesion, LSIL: Low-grade intraepithelial lesion, NILM: Indicates negative for intraepithelial lesions or malignancy—normal or indicative of infection

The assessment of HPV types and colposcopic biopsy outcomes indicated that HPV16, HPV18, non-16/18 HR-HPV, and multiple HPV positivity were not significantly associated with the progression to severe cervical dysplasia (CIN2-3) (respectively, p=0.534, p=0.215, p=0.355, and p=0.285). The relationship between HPV positivity and colposcopic biopsy results is presented in Table 2.

The negative predictive value of cytology for detecting CIN2 or higher lesions (CIN2+) was 92.3%, whereas the positive predictive value was 100%. CIN2+ lesions were absent in 92.3% of patients with negative cytology results (NILM, ASC-US, LSIL), while CIN2+ lesions were detected in all patients with positive cytology findings (HSIL, ASC-H, AGC). Table 3 summarizes the relationship between cytology results and colposcopic biopsy findings.

When patients were categorized by age groups (30-40, 41-50, and 51-65 years), the detection rate of CIN3 was significantly higher in the 51-65 age group (p=0.034). The distribution of colposcopic biopsy outcomes according to age groups is shown in Table 4.

No statistically significant association was observed between menopausal status and colposcopic biopsy results (normal biopsy: p=0.554; CIN1: p=0.601; CIN2: p=0.577; CIN3: p=0.868). Similarly, smoking status was not significantly associated with colposcopic biopsy outcomes (normal biopsy: p=0.509; CIN1: p=0.424; CIN2: p=0.828; CIN3: p=0.433). These findings are summarized in Table 5.

Univariate and multivariate logistic regression analyses were performed to identify factors associated with CIN2+ lesions. None of the evaluated variables, including age, HPV16 positivity, multiple HPV infection, smoking status, and menopausal status, were independently associated with CIN2+ lesions (Table 6).

Table 2. Correlation among HPV subtypes and colposcopy-directed biopsy outcomes

Pathological biopsy results	HPV16 (n=81)	HPV18 (n=29)	Non-16/18 HR-HPV (n=40)	Multiple (dual) HPV positivity (n=30)	p-value*
Normal	54 (66.6%)	21 (72.4%)	32 (80%)	21 (70%)	0.534
CIN1	9 (11.1%)	1 (3.6%)	4 (10%)	1 (3.3%)	0.215
CIN2	3 (3.7%)	2 (6.8%)	1 (2.5%)	3 (10%)	0.355
CIN3	15 (18.6%)	5 (17.2%)	3 (7.5%)	5 (16.7%)	0.285

CIN: Cervical intraepithelial neoplasia, HR-HPV: High-risk human papillomavirus, *: p-values <0.05 were regarded as statistically significant

Table 3. Colposcopy biopsy results according to cytology (smear) results

Pathological biopsy results	NILM (n=127)	ASC-US (n=32)	LSIL (n=9)	HSIL (n=9)	ASC-H (n=2)	AGC (n=1)
Normal	105 (82.6%)	20 (62.5%)	3 (33.3%)	-	-	-
CIN1	8 (6.5%)	4 (12.5%)	3 (33.3%)	-		
CIN2	2 (1.5%)	3 (9.3%)	1 (11.1%)	2 (22.2%)		
CIN3	12 (9.4%)	5 (15.7%)	2 (22.2%)	7 (77.8%)	2 (100%)	1 (100%)

AGC: Abnormal glandular cells, ASC-H: Atypical squamous cells—cannot exclude high-grade squamous intraepithelial lesion, ASC-US: Atypical squamous cells—undetermined significance, CIN: Cervical intraepithelial neoplasia

Table 4. Distribution of colposcopic biopsy outcomes by age categories

Pathological biopsy results	30-40 years (n=80)	41–50 years (n=67)	51-65 years (n=33)	p-value*
Normal	63 (78.7%)	48 (71.6%)	17 (51.5%)	0.214
CIN1	2 (2.5%)	9 (13.4%)	4 (12.1%)	0.391
CIN2	4 (5%)	2 (3.1%)	3 (9.1%)	0.420
CIN3	11 (13.8%)	8 (11.9%)	9 (27.3%)	**0.034**

CIN: Cervical intraepithelial neoplasia, *: p-values <0.05 were regarded as statistically significant

Table 5. Assessment of colposcopy outcomes based on patients' menopausal status and tobacco use			
Pathological biopsy results	**Premenopause (n=136)**	**Postmenopause (n=44)**	**p-value***
Normal	96 (70.5%)	32 (72.7%)	0.554
CIN1	10 (7.3%)	5 (11.4%)	0.601
CIN2	8 (5.9%)	1 (2.2%)	0.577
CIN3	22 (16.3%)	6 (13.7%)	0.868
The effect of smoking on colposcopic biopsy			
Pathological biopsy results	**Non-smoker (n=124)**	**Smoker (n=56)**	**p-value***
Normal	93 (75%)	35 (62.5%)	0.509
CIN1	12 (9.6%)	5 (9%)	0.424
CIN2	6 (4.8%)	3 (5.3%)	0.818
CIN3	15 (12%)	13 (23.2%)	0.433
CIN: Cervical intraepithelial neoplasia, *: p-values <0.05 were regarded as statistically significant			

Table 6. Univariate and multivariate logistic regression analysis for predictors of CIN2+ lesions						
Variable	**Univariate OR**	**95% CI**	**p-value**	**Multivariate OR**	**95% CI**	**p-value**
Age (years)	0.97	0.93-1.01	0.10	0.98	0.94-1.02	0.28
HPV16 positivity	1.68	0.79-3.55	0.18	1.52	0.70-3.31	0.29
Multiple HPV infection	1.94	0.90-4.19	0.09	1.73	0.78-3.84	0.17
Smoking status	1.53	0.77-3.05	0.22	1.41	0.69-2.90	0.34
Postmenopausal status	0.82	0.33-2.03	0.66	0.91	0.36-2.30	0.84
OR: Odds ratio, CI: Confidence interval. Cervical intraepithelial neoplasia (CIN)2+ was defined as cervical intraepithelial neoplasia grade 2 or higher. Variables included in the multivariate model were age, HPV16 positivity, multiple HPV infection, smoking status, and menopausal status, and menopausal status. Univariate and multivariate logistic regression analysis were employed to assess independent variables linked to the outcomes. Statistical significance was defined as p<0.05.						

Discussion

Expertise in anatomy and histology is essential for the accurate performance of clinical colposcopy. Moreover, there are worldwide issues associated with standardizing language, facilitating ongoing training, executing quality assurance methods, and establishing enough infrastructure. According to the results of the study, we determined that our cervical colposcopic evaluation results based on HPV positivity were compatible with the current global and Turkey literature. The lack of adenoCIS and invasive malignancy was a pleasing finding.

The incidence rates of HPV types in our study align with the findings of extensive investigations completed in Turkey (9-11). Despite the absence of a statistically significant difference, current evidence identifies HPV16 as the predominant type associated with cervical dysplasia, particularly CIN2-3, and cervical cancer (12,13). Given that the likelihood of cervical dysplasia and cancer escalates with increased viral load, the literature reveals no statistically significant difference regarding HPV multiple positivity in relation to cervical cancer and cervical dysplasia when compared to other high-risk HPV types, specifically HPV16 and HPV18 (14,15). Furthermore, in contrast to previous claims in the literature, the probability of detecting CIN2 and more severe lesions was not increased in our patients who concurrently tested positive for HPV16-18. As previously shown in the literature, we established the sensitive predictive value of cervical cytology for CIN2 and higher lesions at 92% (16-18). In our study, similar to the study conducted in Turkey, the detection of CIN3 results was found to be high in women aged between 51 and 65 who underwent colposcopic biopsy and endocervical curettage (19).

While our analysis did not demonstrate a statistically significant disparity, we noted that the incidence of CIN3 biopsy findings was elevated in women who smoked and had colposcopic biopsy, consistent with existing literature (20). In accordance with the literature, no significant difference was seen in the identification of CIN2 and higher lesions between premenopausal and postmenopausal individuals when examining colposcopic biopsy data (21).

Study Limitations

The constraints of our study include the limited sample size and its retrospective design. In addition, the strength of our study is that it is a single-center study in which we

determine our own practice.

Conclusion

The findings of our study are generally consistent with both global and Turkish literature. No significant association was observed between HPV genotype distribution and the detection of high-grade cervical lesions (CIN2+) in our cohort. The absence of adenoCIS and invasive cancer in our series may be related to the effectiveness of the national cervical cancer screening program.

Ethics

Ethics Committee Approval: Ethical approval for this study was obtained from the Ethics Committee of University of Health Sciences Turkey, Gaziantep City Hospital (date: 21.05.2025, approval number: 221/2025).

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

Footnotes

Authorship Contributions

Surgical and Medical Practices: E.Ş., G.G., M.G., Concept: E.Ş., G.G., M.G., S.G., Design: E.Ş., M.G., S.G., İ.T., Data Collection or Processing: E.Ş., G.G., M.G., İ.T., Analysis or Interpretation: E.Ş., İ.T., Literature Search:E.Ş., G.G., M.G., S.G., İ.T., Writing: E.Ş.

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

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

References

1. Siegel RL, Kratzer TB, Giaquinto AN, Sung H, Jemal A. Cancer statistics, 2025. CA Cancer J Clin. 2025;75(1):10-45.
2. Pimple S, Mishra G. Cancer cervix: epidemiology and disease burden. Cytojournal. 2022;19:21.
3. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209-249.
4. T.C. Sağlık Bakanlığı. Türkiye kanser istatistikleri 2020. [internet]. Erişim adresi:https://hsgm.saglik.gov.tr/depo/birimler/kanser-db/istatistik/Turkiye_Kanser_istatistikleri_2020.pdf
5. Arbyn M, Weiderpass E, Bruni L, de Sanjosé S, Saraiya M, Ferlay J, et al. Estimates of incidence and mortality of cervical cancer in 2018: a worldwide analysis. Lancet Glob Health. 2020;8(2):e191-e203.
6. Perkins RB, Guido RS, Castle PE, Chelmow D, Einstein MH, Garcia F, et al. 2019 ASCCP risk-based management consensus guidelines for abnormal cervical cancer screening tests and cancer precursors. J Low Genit Tract Dis. 2020;24(2):102-131.
7. Zhou Y, Zhang H, Li X, Chen L, Liu Y. Immediate risk of CIN and diagnostic value of colposcopy among cytology-negative, HR-HPV-positive women. BMC Womens Health. 2024;24:188.
8. Kurtycz DFI, Barkan GA, Savik K, Chhieng DC, Souers RJ, Nayar R, et al. Bethesda interobserver reproducibility study-2 (BIRST-2): Bethesda system 2014. J Am Soc Cytopathol. 2017;6(4):131-144.
9. Gultekin M, Dundar S, Ozgul N, Aydin B, Tuncer ZS, Ayhan A, et al. How to triage HPV-positive cases: results of four million females. Gynecol Oncol. 2020;158(1):105-111.
10. Egemen D, Cheung LC, Chen X, Demarco M, Perkins RB, Kinney W, et al. Risk estimates supporting the 2019 ASCCP risk-based management consensus guidelines. J Low Genit Tract Dis. 2020;24(2):132-143.
11. Gunes AC, Ozgul N, Gultekin M, Akgor U, Turkyilmaz M, Dursun P, et al. Evaluation of colposcopy after the addition of human papillomavirus testing to the Turkish cervical cancer screening program. Cancer Med. 2023;12(24):21751-21760.
12. Priego-Hernández VD, Romero-López C, Muñoz-Bello JO, Rosas-Murrieta NH, Vallejo-Ruiz V, et al. Expression of HIF-1α and genes involved in glucose metabolism is increased in cervical cancer and HPV-16-positive cell lines. Pathogens. 2022;12(1):33.
13. Gori S, Tornesello ML, Buonaguro L, Rossi PG, Boveri S, et al. Clinical relevance of partial HPV16/18 genotyping in stratifying HPV-positive women attending routine cervical cancer screening: a population-based cohort study. BJOG. 2021;128(8):1353-1362.
14. Na J, Li Y, Wang J, Wang X, Lu J, Han S. The correlation between multiple HPV infections and the occurrence, development, and prognosis of cervical cancer. Front Microbiol. 2023;14:1220522.
15. Gökkaya M, Akgul E, Demir O, Ozgul N, Gultekin M. Does HPV-18 co-infection increase the risk of cervical pathology in individuals with HPV-16? Cytopathology. 2024;35(6):757-760.
16. Wu P, Xiong H, Yang M, Li L, Wu P, Lazare C, et al. Co-infections of HPV16/18 with other high-risk HPV types and the risk of cervical carcinogenesis: a large population-based study. Gynecol Oncol. 2019;155(3):436-443.
17. Koyanagi T, Fujiwara H, Yamamoto K, Suzuki M, Kakizoe T. Cervical cancer screening efficacy using SurePath, ThinPrep and conventional cytology: a large data set analysis from the Japan Cancer Society. Cytopathology. 2024;35(6):770-775.
18. Uyar DS, Eltabbakh GH, Mount SL. Positive predictive value of liquid-based and conventional cervical Papanicolaou smears reported as malignant. Gynecol Oncol. 2003;89(2):227-232.
19. Akgor U, Ozgul N, Gunes AC, Turkyilmaz M, Gultekin M. Evaluation of endocervical curettage in colposcopy in the Turkish cervical cancer screening program. J Clin Med. 2024;13(15):4417.
20. Kayar İ, Goc G, Cetin F, Birge Ö. Impact of smoking on cervical histopathological changes in high-risk HPV-positive women: a matched case–control study. Medicina (Kaunas). 2025;61:235.
21. Edwards K, Fatehi M, Fogel J. Post-menopausal status and risk for cervical dysplasia. Gulf J Oncol. 2022;1(38):31-37.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):261-269
DOI: 10.4274/BMB.galenos.2026.70298

The Association of the Wells Score with CT-derived RV/LV Ratio and Mortality

Wells Skorunun BT'den Elde Edilen RV/LV Oranı ve Mortalite ile İlişkisi

Murat Avşar[1], Münevver Gül Avşar[2], Mustafa Hakan Dinçkal[3], İrfan Şahin[4], Süleyman Sezai Yıldız[1]

[1]University of Health Sciences Turkey, Prof. Dr. Cemil Taşcıoğlu City Hospital, Department of Cardiology, İstanbul, Turkey
[2]University of Health Sciences Turkey, Prof. Dr. Cemil Taşcıoğlu City Hospital, Department of Internal Medicine, İstanbul, Turkey
[3]Acıbadem University, Atakent Hospital, Department of Cardiology, İstanbul, Turkey
[4]University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Cardiology, İstanbul, Turkey

Abstract

Objective: Pulmonary thromboembolism (PTE) is the third most common cause of cardiovascular mortality worldwide. The aim of this study was to investigate the relationship between the clinical Wells score and the right ventricle/left ventricle (RV/LV) ratio measured by computed tomography pulmonary angiography (CTPA), and to evaluate the predictive value of these parameters for 30-day mortality in patients diagnosed with PTE.

Method: In this retrospective cross-sectional study, 150 patients diagnosed with acute PTE in the emergency department were included. Wells scores were calculated for all patients, and RV and LV diameters were measured on CTPA images to determine the RV/LV ratio. Parameters associated with mortality were evaluated using univariable and multivariable logistic regression analyses.

Results: Compared with survivors, non-survivors had significantly higher Wells scores (7.06 vs. 4.88; p=0.011), larger RV diameters (45.8 mm vs. 42.4 mm; p=0.034), and higher troponin-I levels (0.52 ng/mL vs. 0.21 ng/mL; p=0.024). Mortality was 23.4% in patients with a Wells score ≥7, whereas no deaths occurred in the low-risk group. Receiver operating characteristic analysis demonstrated an area under the curve of 0.69 for the Wells score and 0.67 for troponin-I in predicting mortality. No significant correlation was found between the Wells score and the CTPA-derived RV/LV ratio (p=0.90). In multivariable analysis, advanced age [odds ratio (OR) 1.13], high Wells score (OR 1.49), and elevated troponin-I levels (OR 2.01) were identified as independent predictors of mortality.

Öz

Amaç: Pulmoner tromboemboli (PTE), dünya genelinde kardiyovasküler mortalitenin üçüncü en sık nedenidir. Bu çalışmanın amacı, PTE tanısı alan hastalarda klinik Wells skoru ile bilgisayarlı tomografi pulmoner anjiyografide (BTPA) ölçülen sağ ventrikül/sol ventrikül (RV/LV) oranı arasındaki ilişkiyi ve bu parametrelerin 30 günlük mortalite üzerindeki öngörücü değerini araştırmaktır.

Yöntem: Retrospektif kesitsel tasarımlı bu çalışmaya, acil serviste akut PTE tanısı alan 150 hasta dahil edilmiştir. Hastaların Wells skorları hesaplanmış, BT görüntülerinden RV ve LV çapları ölçülerek oranlanmıştır. Verilerin analizinde, sürekli değişkenlerin Wells skor grupları arasındaki karşılaştırmaları için tek yönlü varyans analizi kullanılmış; mortalite ile ilişkili parametreler ise univaryant ve multivaryant lojistik regresyon analizleri ile değerlendirilmiştir.

Bulgular: Otuz günlük takipte mortalite oranı %10,6 (n=16) olarak saptanmıştır. Ölen hastalarda Wells skoru (7,06'ya karşı 4,88; p=0,011), sağ ventrikül çapı (45,8 mm'ye karşı 42,4 mm; p=0,034) ve troponin-I düzeyleri (0,52 ng/mL'ye karşı 0,21 ng/mL; p=0,024) sağ kalanlara göre anlamlı derecede yüksektir. Wells skoru ≥7 olan grupta mortalite %23,4 iken, düşük riskli grupta ölüm izlenmemiştir. Alıcı çalışma karakteristiği analizinde Wells skorunun mortaliteyi öngörmede eğri altında kalan alanı 0,69, troponin-I'in ise 0,67 olarak saptanmıştır. Ayrıca malignite varlığı ölen grupta belirgin şekilde fazladır (%25'e karşı %8,3; p=0,048). Wells skoru ile BT pulmoner anjiyografi kaynaklı RV/LV oranı arasında anlamlı bir korelasyon bulunmamıştır (p=0,90). Multivaryant analizde yüksek Wells skoru [olasılık oranı (OR) 1,49],

Address for Correspondence: Murat Avşar, MD, University of Health Sciences Turkey, Prof. Dr. Cemil Taşcıoğlu City Hospital, Department of Cardiology, İstanbul, Turkey
E-mail: muravsar@yahoo.com **ORCID:** orcid.org/0000-0003-0880-8998
Received: 06.01.2026 **Accepted:** 06.06.2026 **Publication Date:** 24.06.2026

Cite this article as: Avşar M, Avşar MG, Dinçkal MH, Şahin İ, Yıldız SS. The association of the Wells score with CT-derived RV/LV ratio and mortality. Bagcilar Med Bull. 2026;11(2):261-269

Abstract

Conclusion: The Wells score and troponin-I levels are critical indicators for identifying increased short-term mortality risk in patients with PTE. Although the Wells score does not show a correlation with the CTPA-derived RV/LV ratio, it is significantly associated with mortality. Troponin-I independently predicts mortality. Our findings suggest that an integrated, multilayered risk stratification approach combining clinical scoring systems and biomarkers may facilitate early identification and optimized management of high-risk PTE patients.

Keywords: Mortality, pulmonary embolism, RV/LV ratio, Wells score

Öz

ileri yaş (OR 1,13) ve troponin-I yüksekliği (OR 2,01) mortaliteyi öngören bağımsız risk faktörleri olarak tanımlanmıştır.

Sonuç: Wells skoru ve troponin-I düzeyi, PTE hastalarında kısa dönem mortalite riskinin artışının belirlenmesinde kritik göstergelerdir. Klinik Wells skoru, BTPA tabanlı RV/LV oranı ile doğrudan korelasyon göstermese de mortalite ile anlamlı ilişkilidir. Troponin-I myokardiyal hasarı yansıtmakla kalmayıp çalışmamızda mortaliteyi de anlamlı derecede öngörmüştür. Çalışmamız klinik skorlama ve biyobelirteçlerin birleştirildiği entegre, çok katmanlı bir risk stratifikasyonu yaklaşımın, yüksek riskli hastaların erken tespiti ve yönetimi için yeni bir ufuk açabileceğini göstermektedir.

Anahtar kelimeler: Mortalite, pulmoner emboli, RV/LV oranı, Wells skoru

Introduction

Pulmonary thromboembolism (PTE) is a life-threatening cardiovascular condition that requires urgent diagnosis and treatment, resulting from the acute obstruction of the main pulmonary artery or its branches by thrombus formation (1,2). PTE is the third most common cause of cardiovascular mortality worldwide, following ischemic heart disease and stroke (3,4). In untreated situation of masive PTE, mortality rates can increase to 25-30% (4,5). The clinical severity of PTE largely depends on the development of acute right ventricular (RV) dysfunction (6,7). In high-risk (massive) PTE cases accompanied by shock, mortality rates can reach 50% (8).

The diagnostic approach to PTE involves clinical pretest probability scoring systems, D-dimer testing, and imaging algorithms (9). The most widely used clinical prediction rules are the Wells score and the revised Geneva score; both assign points based on patients' symptoms, clinical signs, and risk factors to stratify PTE probability into low, intermediate, and high categories (10). The Wells score is the only clinical prediction rule specifically validated in hospitalized patients and was first described by Wells et al. (11), due to its practicality, it is widely used in the initial evaluation of patients with suspected PTE (11). In a 2015 meta-analysis including 11 studies, the Wells score was shown to have superior diagnostic accuracy compared with the revised Geneva score for predicting pulmonary embolism in patients with suspected disease (12). Use of clinical probability assessment tools such as the Wells score allows risk-based application of D-dimer testing and imaging strategies, which may reduce unnecessary computed tomography use (13). These models aim to minimize unnecessary use of computed tomography pulmonary angiography (CTPA) in low-risk patients while ensuring that high-risk PTE cases are not missed (14).

Current guidelines classify PTE into three major risk categories: High, intermediate, and low risk (15). High-risk PTE is defined by the presence of hemodynamic instability, such as hypotension, shock, or cardiac arrest, and represents approximately 5% of cases and is associated with a 30-day mortality of 15-30% (8). In intermediate-high-risk PTE, patients are normotensive but exhibit RV dysfunction and/or positive cardiac injury biomarkers, including troponin and B-type natriuretic peptide (BNP). This subgroup represents approximately 20-25% of all PTE cases and is associated with a 30-day mortality rate of around 5-10%. In contrast, patients with low-risk PTE show no evidence of hemodynamic instability or right ventricular dysfunction (RVD); early mortality in this group is approximately 1%, and outpatient management may be safely considered. (4).

RVD is a key determinant of risk stratification in PTE (15). RVD is defined by an RV/left ventricle (LV) diameter ratio >1.0 on CTPA and/or elevated cardiac biomarkers such as troponin and BNP (16). An increased RV/LV ratio is associated with worsening cardiac output and higher mortality. Patients with RV/LV >1.0 have an approximately fourfold increase in 30-day mortality. In initially stable patients with RVD, mortality may rise substantially if hemodynamic deterioration occurs (17,18). Therefore, early recognition of RV overload is essential to guide the timely escalation of treatment in PT.

While the Wells score has been primarily validated and widely used as a diagnostic tool for assessing the probability of PTE, its role as a prognostic indicator, particularly in relation to RVD and short-term mortality, remains less well established. This study, therefore, aims not only to examine

the diagnostic utility of the Wells score but also to evaluate its potential prognostic value in predicting mortality and its correlation with imaging markers of RV involvement in patients with acute PTE.

Materials and Methods

This single-center study was a retrospective cross-sectional analysis. We reviewed the medical records of patients aged ≥18 years who were diagnosed with acute PTE and admitted to the emergency department before 2018. Patients were eligible if pulmonary artery thromboembolism was confirmed by CTPA and complete clinical and laboratory data were available from archival records. Patients were excluded for chronic pulmonary hypertension, pregnancy, trauma recent major surgery requiring prophylactic anticoagulation, concomitant severe cardiac disease such as advanced left heart failure or incomplete data. The study protocol was approved by the Ethics Committee of Bezmialem Vakıf University (approval number: E-540224451-050.04-216335, date: 31.10.2025).

Demographic characteristics (age, sex, and comorbidities), clinical findings at presentation, vital signs (blood pressure, heart rate, and oxygen saturation), and laboratory data were obtained through a review of patient medical records and digital archive systems. Venous blood samples were collected from the antecubital vein using vacuum tubes containing EDTA for complete blood count, serum-separator tubes for biochemical and cardiac biomarker analyses, and 3.2% citrate tubes for D-dimer measurement. Biochemical parameters were analyzed using the Siemens Diagnostics ADVIA 1800 system. Cardiac troponin I was measured with the Siemens Diagnostics ADVIA Centaur XP immunoassay analyzer, and D-dimer levels were determined using an enzyme-linked immunosorbent assay (VIDAS ELISA).

For each patient, the Wells clinical prediction score was calculated using data obtained at presentation. The score components included clinical signs of deep vein thrombosis (3 points); tachycardia, defined as a heart rate >100 beats/min (1.5 points); immobilization or recent major surgery (1.5 points); history of PTE or deep vein thrombosis (1.5 points); hemoptysis (1 point); active malignancy (1 point); and absence of an alternative diagnosis more likely than PTE (3 points). Based on the total score, patients were classified into three probability categories: Low (0-1 points), intermediate (2-6 points), and high (≥7 points).

CTPA images obtained at the time of diagnosis were retrospectively reviewed by two independent radiologists who were blinded to the clinical data. RV and LV diameters were measured on axial images representing the four-chamber view, from the inner wall to the inner wall. Measurements were preferentially performed during the diastolic phase of the cardiac cycle, when ventricular dimensions are maximal, and the largest diameters were recorded in millimeters. The RV/LV diameter ratio was calculated for each patient. In accordance with previously established thresholds, the presence of RV dilatation on CTPA was assessed using an RV/LV ratio ≥1.0. (18-20).

CTPA reports were reviewed to determine localization of thrombus within the pulmonary arterial tree and to document ancillary findings, including cardiac chamber enlargement, interventricular septal deviation, and inferior vena cava distension. When available, echocardiographic findings at initial presentation and cardiac biomarker levels (troponin I and BNP) were retrieved from archival records and recorded as markers of RVD.

The echocardiographic data were evaluated retrospectively. Records of transthoracic echocardiography (TTE) examinations, performed as part of routine clinical care within the first 24 hours after hospital admission, were analyzed. Examinations were performed using a cardiac ultrasound system (Epiq 7 Philips Ultrasound System, Amsterdam, the Netherlands) in accordance with the recommendations of the American Society of Echocardiography (ASE). RV dimension was assessed by measuring the basal RV diameter from the apical four-chamber view. RV systolic function was evaluated using tricuspid annular plane systolic excursion (TAPSE). Left ventricular ejection fraction (LVEF) was calculated using the biplane Simpson method from apical two- and four-chamber views. For analyses involving biomarkers and echocardiographic parameters, only patients with available data for the respective variables were included. No imputation was performed due to the retrospective nature of the study.

Statistical Analysis

Statistical analyses were performed using SPSS software (version 22.0). Continuous variables were presented as mean ± standard deviation or median (interquartile range) depending on data distribution, and categorical variables were expressed as counts and percentages. Normality of continuous data was assessed using the Shapiro-Wilk test.

Comparisons between two groups were performed using the independent samples t-test or Mann-Whitney U test, as appropriate. For comparisons among more than two groups, One-Way ANOVA or Kruskal-Wallis tests were used. Categorical variables were compared using the chi-square test or Fisher's exact test, as appropriate. Correlations between continuous variables were evaluated using Spearman's rank correlation coefficient. Receiver operating characteristic (ROC) curve analyses were performed to determine the predictive ability of the Wells score and troponin-I for 30-day mortality, and the area under the curve (AUC), sensitivity, and specificity were calculated. Univariable logistic regression analyses were conducted to identify variables associated with mortality; those with p<0.05 were subsequently included in a multivariable logistic regression model to determine independent predictors. A two-sided p-value <0.05 was considered statistically significant.

Results

A total of 150 patients were included in the study. During the 30-day follow-up period, 16 patients (10.6%) died. The clinical, demographic, laboratory, and radiological characteristics of the patients are summarized in Table 1. The mean age of the study population was 63.5±16.1 years.

Compared with survivors, patients in the mortality group were significantly older (77 vs. 61 years, p<0.001) and had a higher proportion of female patients (54%). In addition, the mortality group had a significantly larger RV end-diastolic diameter (45.8 mm vs. 42.4 mm, p=0.034). Troponin-I levels were also significantly higher in non-survivors than in survivors (0.52 ng/mL vs. 0.21 ng/mL, p=0.024). Similarly, the RV/LV ratio was higher in the mortality group. Moreover, the Wells score was significantly higher in patients who died (7.06 vs. 4.88, p=0.011), and the prevalence of malignancy was greater among patients who died (25% vs.

Table 1. Distribution of demographic characteristics, vital signs, laboratory findings, imaging parameters, and comorbidity prevalence in the overall study population, survivors, and non-survivors

Parameters	All patients (n=150)	Survive (n=134)	Non-survive (n=16)	p-value
Age	63.5	61.85	77.31	**<0.001**
Female	68	61	7	0.92
Male	82	73	9	0.92
Height (cm)	166.77	166.89	166.06	0.826
Weight (kg)	78.4	79.11	74.0	0.306
BMI	26.88	27.36	24.73	0.072
Pulse	92.9	91.3	104.38	0.054
Temperature	36.69	36.69	36.68	1
Respiratory rate	22.12	21.97	23.12	0.846
Troponin-I	0.25	0.21	0.52	0.024
LV (mm) measured by CT	47.38	47.31	47.98	0.695
RV (mm) measured by CT	42.78	42.41	45.84	0.034
RV/LV measured by CT	0.92	0.91	0.99	0.023
Wells score	5.15	4.88	7.06	**0.011**
DM %	27.6	25.4	43.8	0.142
HT %	45.6	47.5	31.2	0.289
CAD %	11.9	11.9	12.5	1
CKD %	6.0	5.1	12.5	0.244
CVD %	8.2	6.8	18.8	0.127
CHF %	12.6	12.6	12.5	1
Malignancy %	10.3	8.3	25.0	**0.048**
RVd (mm) measured by ECHO	36.8	35.9	42.1	**<0.001**
TAPSE (mm)	18.2	18.9	14.6	**<0.001**
LVEF (%)	57.4	57.8	55.9	0.21

Data are presented as mean ± standard deviation or n (percentage). Statistically significant p-values are shown in bold. BMI: Body mass index, LV: Left ventricle, RV: Right ventricle, RV/LV: Right ventricle/left ventricle diameter ratio, DM: Diabetes mellitus, HT: Hypertension, CAD: Coronary artery disease, CKD: Chronic kidney disease, CVA: Cerebrovascular accident, CHF: Congestive heart failure, LVEF: Left ventricular ejection fraction, TAPSE: Tricuspid annular plane systolic excursion

8.3%, p=0.048). On echocardiographic assessment, non-survivors had a significantly larger RV basal diameter (42.1 mm vs. 35.9 mm, p<0.001) and lower TAPSE values (14.6 mm vs. 18.9 mm, p<0.001), while no significant difference was observed in LVEF (55.9% vs. 57.8%, p=0.21).

In the overall cohort, 27.6% of patients had diabetes mellitus, 45.6% had hypertension, 27.7% were current smokers, 32.3% had a history of immobilization, and 36% had deep vein thrombosis.

When patients were stratified into low-, intermediate-, and high-risk groups according to the Wells score, the demographic, clinical, laboratory, and radiological parameters were summarized in Table 2. No significant differences were observed among the groups in age, height, body mass index, systolic and diastolic blood pressure, troponin-I levels, or RV/LV ratio measured on CTPA (all p>0.05). The mean Wells score for the entire cohort was 5.1, whereas it was 7.1 in patients who died and 4.9 in survivors.

Mortality increased significantly with increasing Wells score. No deaths occurred in the low-risk group, whereas mortality was 7.2% in the intermediate-risk (3-6) group and 23.4% in the high-risk (≥7) group. This difference was statistically significant (p=0.011) (Figure 1).

Using a Wells score cut-off value of ≥7, the sensitivity and specificity for predicting 30-day mortality in patients with PTE were 68.8% and 68.4%, respectively. ROC curve analysis demonstrated an AUC of 0.69 [95% confidence interval (CI): 0.56-0.81; p=0.006] (Figure 2a). The cut-off value for troponin-I was set at 0.014 ng/mL, and 34.7% of the patients had troponin levels >0.014 ng/mL. The mean troponin level was 0.52 ng/mL in patients who died and 0.26 ng/mL in survivors (p=0.024). In addition, the mortality rate was 31.0% among patients with positive troponin levels, compared with 14.9% in those with normal troponin levels. Troponin positivity significantly predicted mortality (p=0.024). A ROC analysis was performed to evaluate the association between elevated troponin levels and mortality. Using a troponin-I cut-off value of 0.014 ng/mL, the analysis showed that troponin-I demonstrated moderate discriminative ability for predicting mortality (AUC =0.67) (95% CI: 0.33-0.80; p=0.023). The sensitivity was 87.5% and the specificity was 43.3% (Figure 2b).

Table 2. Distribution of demographic characteristics, vital signs, comorbidities, and mortality among patients with low, intermediate, and high Wells scores

Parameters	Low (≤2) (n=14)	Intermediate (3-6) (n=69)	High (≥7) (n=47)	p-value
Age	63.93±15.22	64.80±15.45	63.77±17.08	0.9392
Height	165.50±6.91	165.79±7.65	168.85±9.19	0.1561
Weight	79.00±13.54	77.13±13.23	79.95±18.08	0.6479
BMI	27.99±2.33	26.84±4.90	26.64±4.27	0.7712
BP systolic	120.71±31.68	117.35±19.15	112.00±19.49	0.2724
BP diastolic	77.79±25.98	72.24±11.83	69.16±12.65	0.1343
Pulce	99.14±24.87	88.84±16.21	95.51±22.23	0.0837
Temperature	36.60±0.32	36.70±0.54	36.68±0.44	0.7747
Respiratory rate	20.86±2.18	22.06±4.24	22.51±4.61	0.4405
Troponin-I	0.13±0.22	0.38±1.03	0.14±0.18	0.2378
RV/LV	0.89±0.16	0.93±0.22	0.92±0.20	0.7620
DM %	35.7	21.7	36.2	0.1955
HT %	57.1	40.6	51.1	0.3681
CAD %	28.6	13.0	6.4	0.0823
CKD %	14.	5.8	4.3	0.3845
CVD %	0.	5.8	14.9	0.1088
CHF %	85.7	4.3	2.1	<0.0001
Malignancy %	0.	8.7	10.6	0.4524
COPD %	35.7	8.7	8.5	0.0112
Mortality %	0.0 (0/14)	7.2 (5/69)	23. (11/47)	0.0111

BMI: Body mass index, BP: Blood pressure, RV/LV: Right ventricle /left ventricle diameter ratio, DM: Diabetes mellitus, HT: Hypertension, CAD: Coronary artery disease, CKD, Chronic kidney disease, CVA: Cerebrovascular accident, CHF: Congestive heart failure, COPD: Chronic obstructive pulmonary disease.

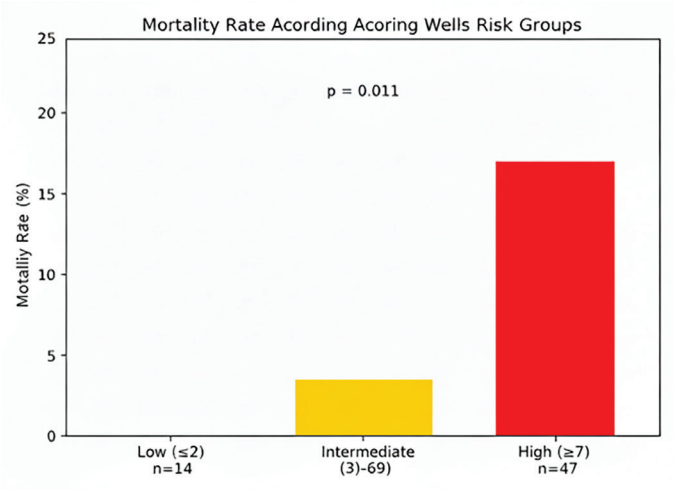


Figure 1. Mortality rate according to Wells risk groups

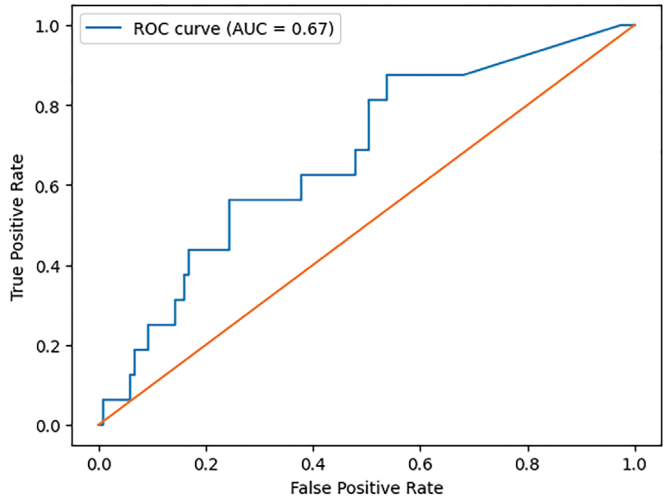


Figure 2b. ROC analysis demonstrating the association between troponin-I and mortality

ROC: Receiver operating characteristic, AUC: Area under the curve

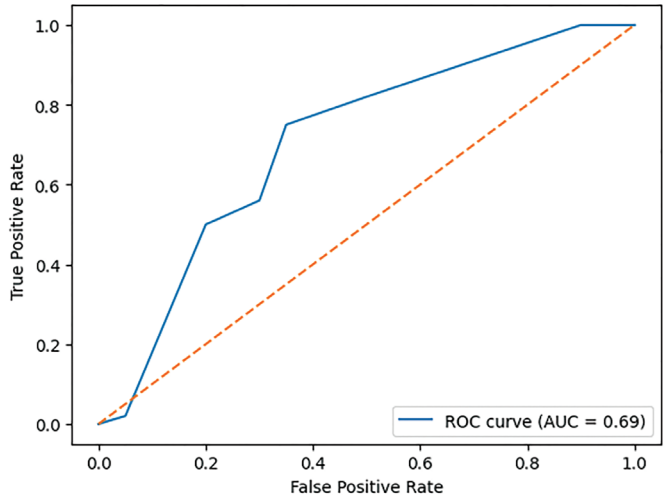


Figure 2a. ROC curve of the Wells score for mortality

ROC: Receiver operating characteristic, AUC: Area under the curve

The RV/LV ratio assessed by CTPA was greater than 0.9 in 48% of the patients. No assessed was found between the Wells score and the CTPA-derived RV/LV ratio (Spearman's ρ=-0.01; p=0.90). No significant difference in Wells scores between patients with and without RV dilatation was observed (p=0.33). An RV/LV ratio >1 was detected in 24% of the patients. The mortality rate in this group was 28.6%, whereas it was 16.4% in patients with an RV/LV ratio ≤1.

In univariable logistic regression analysis, age, Wells score, and troponin-I level were significantly associated with mortality, whereas the RV/LV ratio was not. Variables that were significant in the univariable analysis were subsequently included in the multivariable logistic regression model. In the multivariable analysis, age [odds ratio (OR) 1.13; 95% CI, 1.05-1.21; p<0.001], Wells score (OR 1.49; 95% CI, 1.16-1.92; p=0.002), and troponin-I level

Table 3. Univariable and multivariable logistic regression analyses

Variable	Univariable OR (95% CI)	p-value	Multivariable OR (95% CI)	p-value
Years	1.12 (1.06-1.18)	<0.001	1.13 (1.05-1.21)	<0.001
Wells score	1.42 (1.18-1.71)	<0.001	1.49 (1.16-1.92)	0.002
Troponin-I	1.85 (1.10-3.12)	0.021	2.01 (1.03-3.94)	0.041
Malignancy	2.84 (1.01-7.96)	0.048	30.9 (2.80-341.3)	0.005
cystolic BP ≤90 mmHg	3.92 (1.18-13.01)	0.026	2.51 (0.32-19.84)	0.384

Variables associated with mortality were initially evaluated using univariable logistic regression analysis and subsequently using multivariable logistic regression analysis. Results are presented as odds ratios (ORs), 95% confidence intervals (CIs), and p-values. BP: Blood pressure

(OR 2.01; 95% CI, 1.03-3.94; p=0.041) were independently associated with mortality (Table 3).

Discussion

In the present study, the Wells score, elevated troponin levels, and advanced age were significantly associated with mortality in patients with acute PTE. Notably, although higher troponin-I levels were associated with increased mortality and higher RV/LV ratios, there was no significant correlation between the Wells score and RV/LV ratios.

The RV/LV ratio has been widely investigated as an imaging marker of RVD and adverse outcomes in PTE. Previous studies have demonstrated that an increased RV/LV ratio is associated with reduced cardiac index, particularly in intermediate-risk patients, and may predict hemodynamic deterioration (19). An emergency-department-based cohort reported that the RV/LV ratio had high sensitivity for predicting 30-day mortality (20). D espite these findings, the clinical utility of the RV/LV ratio remains limited by substantial heterogeneity in measurement techniques and cut-off values. A comprehensive review, encompassing data from 35 studies, highlighted considerable variability in RV/LV ratio assessment methods and supported the use of an RV/LV ratio >1.0 measured on a single axial slice as a practical and reproducible approach. In our cohort, the absence of a correlation between the Wells score and the RV/LV ratio may be explained by this lack of standardization, as well as by the sensitivity of RV/LV ratio measurements to slice orientation, cardiac cycle phase, and operator-dependent variability (21). These factors may attenuate statistical associations, particularly in single-center studies with heterogeneous risk profiles.

In contrast, troponin-I emerged as a robust prognostic marker in our study. Elevated troponin levels were independently associated with early mortality among patients with normotensive pulmonary embolism. Mean troponin levels were substantially higher in non-survivors than in survivors. These findings are consistent with contemporary literature confirming the independent association between troponin elevation and short-term mortality in pulmonary embolism.

Troponin elevation reflects RV pressure overload, ischemia, and myocardial injury, and therefore serves as a direct marker of the severity of RVD in pulmonary embolism. In our study, elevated troponin levels were associated with a substantially higher mortality risk compared with normal levels (22). This finding supports the concept that cardiac biomarkers provide incremental prognostic information beyond clinical risk scores alone and, in certain clinical settings, may offer greater prognostic utility than imaging-based parameters such as the RV/LV ratio. The concomitant presence of elevated troponin levels and an increased RV/LV ratio appears to define a higher-risk clinical profile, potentially necessitating closer clinical monitoring and more intensive management strategies. In contrast, the combination of normal troponin levels and an RV/LV ratio ≤1 may identify a relatively low-risk subgroup of patients, in whom outpatient management could be considered in carefully selected cases (23).

Accumulating evidence supports the use of multimodal risk stratification strategies that integrate clinical risk scores, imaging findings, and cardiac biomarkers. In particular, the pulmonary embolism severity index (PESI) and its simplified version, which incorporate variables such as age, comorbidities, and hemodynamic parameters, have been extensively validated as independent predictors of short- and long-term mortality in patients with pulmonary embolism. Several recent studies have shown that combining clinical scores (such as Wells, Geneva, and PESI) with imaging markers like the RV/LV ratio and biomarkers provides superior prognostic discrimination compared with reliance on any single parameter alone (24). Consistent with this framework, our findings suggest that patients with both low clinical risk and limited imaging evidence of RVinvolvement may represent a subgroup suitable for less-intensive management, with potential implications for safe outpatient care and more efficient use of healthcare resources.

In the context of current clinical practice, our findings can be interpreted alongside the ESC/ERS risk stratification framework for acute PTE. According to these guidelines, patients are classified into low, intermediate-low, intermediate-high, and high-risk categories based on hemodynamic status RVD, and cardiac biomarker elevation. Our results—demonstrating that elevated troponin levels and increased RV/LV ratios are associated with higher mortality, while low Wells scores and normal imaging parameters may indicate lower risk— align with this stratification system. These observations support the clinical relevance of combining easily obtainable clinical scores, imaging findings, and biomarkers to refine risk assessment and guide management decisions, including the selection of patients suitable for outpatient care or closer monitoring.

To our knowledge, this study is among the first to directly examine the relationship between the Wells score, RV/LV ratio, and mortality in patients with PTE. While previous studies have largely addressed the diagnostic performance of the Wells score or the prognostic value of imaging markers in isolation, our findings suggest that the Wells score may also have prognostic relevance (25). These results highlight that a simple and widely accessible clinical score can provide meaningful prognostic information, potentially complementing imaging- and biomarker-based risk stratification without requiring advanced diagnostic modalities.

Advanced age was strongly associated with mortality in our cohort, consistent with current European Society of Cardiology/European Respiratory Society guidelines, which recognize age as an independent predictor of adverse outcomes in PTE. Taken together, our findings emphasize the value of an integrated risk stratification approach that combines clinical assessment, cardiac biomarkers, and imaging parameters. In this context, the coexistence of troponin elevation and an increased RV/LV ratio appears to identify a higher-risk patient profile that may warrant closer clinical surveillance (26).

Study Limitations

This study has several limitations. Its retrospective and single-center design may limit the generalizability of the findings. The relatively small sample size may have reduced the statistical power of the analyses. Potential heterogeneity in RV/LV ratio measurement protocols on CTPA and the lack of standardized measurement techniques may have introduced variability and bias. A further limitation of this study is that cardiac troponin levels were measured at a single time point; the absence of serial measurements may have limited our ability to evaluate dynamic changes and temporal trends in troponin levels. Another limitation of this study is that interobserver variability in RV and LV measurements was not formally assessed, which may have affected measurement reliability. Therefore, the results should be interpreted with caution and validated in prospective studies with larger sample sizes and standardized imaging protocols.

Conclusion

Although no significant association was observed between the Wells score and the RV/LV ratio, concordance was observed between the Wells score and troponin levels. These findings may suggest that clinical risk scores are more closely aligned with biomarker-defined myocardial injury than with imaging-based measures alone. Overall, our results support the potential value of an integrated, multilayered risk stratification approach combining clinical assessment, biomarkers, and imaging for short-term prognostic evaluation in patients with pulmonary embolism. Larger prospective studies are warranted to further clarify and confirm these observations.

Ethics

Ethics Committee Approval: The study protocol was approved by the Ethics Committee of Bezmialem Vakıf University (approval number: E-540224451-050.04-216335, date: 31.10.2025).

Informed Consent: Retrospective study.

Footnotes

Authorship Contributions

Surgical and Medical Practices: M.A., M.H.D., İ.Ş., S.S.Y., Concept: M.A., M.G.A., M.H.D., İ.Ş., S.S.Y., Design: M.A., M.G.A., M.H.D., S.S.Y., Data Collection or Processing: M.A., M.G.A., M.H.D., İ.Ş., S.S.Y., Analysis or Interpretation: M.A., M.G.A., M.H.D., İ.Ş., S.S.Y., Literature Search: M.A., M.G.A., İ.Ş., S.S.Y., Writing: M.A., M.G.A., İ.Ş., S.S.Y.

Conflict of Interest: One of the authors of this article (İ.Ş.) is a member of the Editorial Board of this journal. He was completely blinded to the peer review process of the article. No conflict of interest was declared by the authors.

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

References

1. Hecker M, Bayer M, Krombach GA, Sommer N. Diagnostik und therapie der akuten lungenembolie [Diagnosis and treatment of acute pulmonary embolism]. Med Klin Intensivmed Notfmed. 2025;120(3):256-269. German.
2. Latsios G, Mantzouranis E, Kachrimanidis I, Theofilis P, Dardas S, Stroumpouli E, et al. Recent advances in risk stratification and treatment of acute pulmonary embolism. World J Cardiol. 2025;17(5):104983.
3. Fumarola F, Colombo F, Corsaro C, Savi L, Mangione R, Savio D, et al. Role of transcatheter treatment in intermediate-to high-risk pulmonary Embolism. Interv Cardiol. 2025;20:e17.
4. Walker Y, Chiou JP, Diaz-Miret M. Case report of pulmonary embolism with right ventricular strain in a young female. Arch Med Case Rep. 2023;5(1):34-37.
5. Katyal A, Joshi A. Managing right ventricular failure in the setting of pulmonary embolism. Cleve Clin J Med. 2025;92(7):396.
6. Vyas V, Sankari A, Goyal A. Acute pulmonary embolism. [Updated 2024 Dec 11]. In: StatPearls [Internet]. Treasure Island (FL):

StatPearls Publishing; 2024. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560551/

7. Burke M, Khandhar SJ. Managing right ventricular failure in pulmonary embolism. Cardiac Interv Today. 2023;17-21.
8. Peracaula M, Sebastian L, Francisco I, Vilaplana MB, Rodríguez-Chiaradía DA, Tura-Ceide O. Decoding pulmonary embolism: pathophysiology, diagnosis, and treatment. Biomedicines. 2024;12(9):1936.
9. Thomas SE, Weinberg I, Schainfeld RM, Rosenfield K, Parmar GM. Diagnosis of pulmonary embolism: a review of evidence-based approaches. J Clin Med. 2024;13(13):3722.
10. Klok FA, Kruisman E, Spaan J, Nijkeuter M, Righini M, Aujesky D, et al. Comparison of the revised Geneva score with the Wells rule for assessing clinical probability of pulmonary embolism. J Thromb Haemost. 2008;6(1):40-44.
11. Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, et al. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and D-dimer. Ann Intern Med. 2001;135(2):98-107.
12. Shen JH, Chen HL, Chen JR, Xing JL, Gu P, Zhu BF. Comparison of the Wells score with the revised Geneva score for assessing suspected pulmonary embolism: a systematic review and meta-analysis. J Thromb Thrombolysis. 2016;41(3):482-492.
13. Tang L, Hu Y, Pan D, Yang C, Tang C, Huang Y, et al. PECSS: pulmonary embolism comprehensive screening score to safely rule out pulmonary embolism among suspected patients presenting to emergency department. BMC Pulm Med. 2023;23(1):287.
14. Glazier MM, Glazier JJ. Diagnostic strategies in pulmonary embolism. Int J Angiol. 2024;33(2):89-94.
15. Roy PM, Sanchez O, Huisman M, Jimenez D. Risk stratification of acute pulmonary embolism. Presse Med. 2024;53(3):104243.
16. Henzler T, Roeger S, Meyer M, Schoepf UJ, Nance JW Jr, Haghi D, et al. Pulmonary embolism: CT signs and cardiac biomarkers for predicting right ventricular dysfunction. Eur Respir J. 2012;39(4):919-926.
17. Eguchi S, Orihara Y, Eguchi A, Pfeiffer M, Peterson B, Ruzieh M, et al. Additive value of right ventricular global longitudinal strain to a conventional echocardiographic parameter to improve prognostic value in intermediate-risk pulmonary embolism. J Am Heart Assoc. 2025;14(7):e036294.
18. Cho SU, Cho YD, Choi SH, Yoon YH, Park JH, Park SJ, et al. Assessing the severity of pulmonary embolism among patients in the emergency department: utility of RV/LV diameter ratio. PLoS One. 2020;15(11):e0242340.
19. İşcan Er B, Melekoğlu A, Altınbilek E, Öztürk D. Relationship between shock index, sPESI score, and right ventricular dysfunction in CTPA with mortality in patients diagnosed with acute PTE in the emergency department. J Acad Res Med. 2023;13(3):131-137.
20. Meinel FG, Nance JW Jr, Schoepf UJ, Hoffmann VS, Thierfelder KM, Costello P, et al. Predictive value of computed tomography in acute pulmonary embolism: systematic review and meta-analysis. Am J Med. 2015;128(7):747-759.e2.
21. Hanmandlu A, Mannari A, Abidov A. Diagnostic and prognostic value of RV strain defined as RV/LV ratio on chest CT in acute pulmonary embolism: a systematic review. Echocardiography. 2025;42(3):e70107.
22. Yang L, Li B, Chen H, Belfeki N, Monchi M, Moini C. The role of troponin in the diagnosis and treatment of acute pulmonary embolism: mechanisms of elevation, prognostic evaluation, and clinical decision-making. Cureus. 2024;16(8):e67922.
23. Ebner M, Guddat N, Keller K, Merten MC, Lerchbaumer MH, Hasenfuß G, et al. High-sensitivity troponin I for risk stratification in normotensive pulmonary embolism. ERJ Open Res. 2020;6(4):00625-2020.
24. Jiménez D, Bikdeli B, Rodríguez C, Muriel A, Ballaz A, Soler S, et al; RIETE investigators. Identification of low-risk patients with acute symptomatic pulmonary embolism. Arch Bronconeumol. 2023;59(9):575-580.
25. Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, et al. 2019 ESC guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). Eur Heart J. 2020;41(4):543-603.
26. Nishikawa R, Yamashita Y, Morimoto T, Kaneda K, Chatani R, Nishimoto Y, et al. Selection of home treatment and identification of low-risk patients with pulmonary embolism based on simplified pulmonary embolism severity index score in the era of direct oral anticoagulants. J Am Heart Assoc. 2024;13(19):e034953.

ORIGINAL RESEARCH

Bagcilar Med Bull 2026;11(2):270-278
DOI: 10.4274/BMB.galenos.2026.88700

Uric Acid-to-albumin Ratio as an Independent Predictor of Short-term Mortality in Acute Pulmonary Embolism

Akut Pulmoner Embolide Kısa Dönem Mortalitenin Bağımsız Bir Belirteci Olarak Ürik Asit/Albümin Oranı

Orhan İnce[1], Esra Dönmez[2], Muhammed Furkan Deniz[1], Cemal Terzioğlu[1], Sercan Bulut[1], Sevgi Özcan[1]

[1]University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Cardiology, İstanbul, Turkey
[2]University of Health Sciences Turkey, Başakşehir Çam and Sakura City Hospital, Department of Cardiology, İstanbul, Turkey

Abstract

Objective: We aimed to evaluate whether the uric acid-to-albumin ratio (UAR) predicts 30-day mortality in patients with acute pulmonary embolism (PE).

Method: This retrospective study included adult patients diagnosed with acute PE between June 2011 and August 2024. After applying exclusion criteria, 261 patients were analyzed. PE diagnosis was confirmed by computed tomographic pulmonary angiography. The primary endpoint was 30-day all-cause mortality. ROC curve analysis was performed to determine the optimal UAR cut-off value. Independent predictors of mortality were identified using multivariable logistic regression analysis.

Results: The mean age of the study population was 63.6±15.9 years, and 60.2% were female. During the 30-day follow-up, 49 patients (18.8%) died. Non-survivors were significantly older. Furthermore, these patients exhibited significantly higher values for the UAR, systolic pulmonary artery pressure, and PE severity index scores. Also, a statistically significant increase in the frequency of a right ventricle/left ventricle ratio exceeding 1 and a history of malignancy, as well as significantly lower values albumin and the estimated glomerular filtration rate, were detected in non-survivors. ROC analysis specified a UAR cut-off value of 1.78 for predicting 30-day mortality [area under the curve: 0.788; 95% confidence interval (CI): 0.716-0.860], with 71.4% sensitivity and 67.9% specificity. In multivariable logistic

Öz

Amaç: Bu çalışmada, akut pulmoner emboli (PE) hastalarında ürik asit/albümin oranının (ÜAO) 30 günlük mortaliteyi öngörmedeki değerini araştırmayı amaçladık.

Yöntem: Haziran 2011-Ağustos 2024 tarihleri arasında akut PE tanısı alan 261 hasta retrospektif olarak değerlendirildi. PE tanısı bilgisayarlı tomografi pulmoner anjiyografi ile doğrulandı. Birincil sonlanım noktası 30 günlük tüm nedenlere bağlı mortalite olarak belirlendi. ÜAO için optimal kesim değeri ROC eğrisi analizi ile belirlendi. Mortalite için bağımsız prediktörler çok değişkenli lojistik regresyon analizi ile değerlendirildi.

Bulgular: Çalışma popülasyonunun yaş ortalaması 63,6±15,9 yıl olup, hastaların %60,2'si kadındı. Otuz günlük mortalite 49 hastada (%18,8) saptandı. Mortalite izlenen hastalar anlamlı olarak daha ileri yaşta olup ÜAO, sistolik pulmoner arter basıncı, PE şiddet indeks skorları, sağ ventrikül/sol ventrikül oranı 1'den büyük olma ve malignite öyküsü sıklığı daha fazla saptandı. Bunun yanı sıra tahmini glomerüler filtrasyon hızı mortalite izlenen grupta belirgin olarak daha düşük bulundu. ROC analizinde ÜAO için 1,78 kesim değeri 30 günlük mortaliteyi %71,4 duyarlılık ve %67,9 özgüllük ile öngördü [eğri altında kalan alan: 0,788; %95 güven aralığı (GA): 0,716-0,860]. Çok değişkenli lojistik regresyon analizinde yüksek ÜAO [olasılık oranı (OO): 4,798; %95 GA: 2,485-9,262; p<0,001] ve malignite öyküsü (OO: 3,343; %95 GA: 1,218-9,170; p=0,019) 30 günlük mortalitenin bağımsız prediktörleri olarak saptandı.

Address for Correspondence: Orhan İnce, MD, University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital, Department of Cardiology, İstanbul, Turkey
E-mail: drorhanince@gmail.com **ORCID:** orcid.org/0000-0001-6413-6407
Received: 18.02.2026 **Accepted:** 06.06.2026 **Epub:** 10.06.2026 **Publication Date:** 24.06.2026

Cite this article as: İnce O, Dönmez E, Deniz MF, Terzioğlu C, Bulut S, Özcan S. Uric acid-to-albumin ratio as an independent predictor of short-term mortality in acute pulmonary embolism. Bagcilar Med Bull. 2026;11(2):270-278

Abstract

regression analysis, high UAR [odds ratio (OR) 4.798; 95% CI 2.485-9.262; p<0.001] and history of malignancy (OR 3.343; 95% CI 1.218-9.170; p=0.019) were independent predictors of 30-day mortality.

Conclusion: In acute PE, UAR was identified as an independent risk factor for 30-day mortality. This inexpensive biomarker, obtained from routine tests, can contribute to early risk stratification and aid in determining treatment strategies when integrated into existing prognostic models. However, large-scale, multicenter prospective studies are needed to confirm these results and demonstrate the additional clinical benefit of UAR in assessing the prognosis of acute PE patients.

Keywords: Acute pulmonary embolism, mortality, prognosis, risk stratification, uric acid-to-albumin ratio

Öz

Sonuç: Akut PE hastalarında ÜAO'ı 30 günlük mortalitenin bağımsız bir risk belirteci olarak saptanmıştır. Rutin laboratuvar testlerinden elde edilen bu düşük maliyetli biyobelirteç, mevcut prognostik modellere entegre edildiğinde erken risk sınıflamasına katkı sağlayabilir ve daha yoğun tedavi stratejilerinden fayda görebilecek hastaların belirlenmesine katkı sağlayabilir. Ancak bu bulguların doğrulanması ve ÜAO'ın ek klinik katkısının ortaya konulması için geniş ölçekli, çok merkezli ve prospektif çalışmalara ihtiyaç vardır.

Anahtar kelimeler: Akut pulmoner emboli, mortalite, prognoz, risk sınıflaması, ürik asit/albümin oranı

Introduction

The acute pulmonary embolism (PE) in-hospital mortality rate ranges between 5% and 10%. Among cardiovascular diseases, it constitutes the third most common cause of mortality, following myocardial infarction and stroke (1,2). Approximately 5% of patients with PE present with hemodynamic instability or cardiopulmonary arrest secondary to acute right ventricular failure, with a mortality rate of about 30% (3). Therefore, risk stratification and prognosis assessment are crucial for determining the most appropriate treatment strategy.

Patients with hemodynamic instability are classified as high-risk. Most patients do not have hemodynamic instability. These patients require further risk assessment by using clinical conditions, comorbidities, and laboratory and imaging indicators of PE severity, such as cardiac troponins, heart-type fatty acid-binding protein (H-FABP), N-terminal pro-brain natriuretic peptide (NT-proBNP) levels, and signs of right ventricular dysfunction detected by echocardiography or computed tomographic pulmonary angiography (CTPA). Hence, different combinations of these variables have been employed to develop 30-day mortality prediction scores. Among clinical scores that combine PE severity with underlying disease, the pulmonary embolism severity index (PESI) currently possesses the most validation evidence (4). These scoring systems provide valuable prognostic information regarding short-term mortality risk but cannot precisely predict mortality for every individual patient. So, additional factors are needed to increase the predictive value of mortality in acute PE patients (5).

Uric acid (UA) promotes hypercoagulability through oxidative stress, inflammation, and endothelial dysfunction. During its production via the xanthine oxidase pathway, reactive oxygen species are formed, which reduces the protective nitric oxide level and causes endothelial damage. Elevated serum UA level is associated with endothelial activation and increased production of inflammatory cytokines, which enhance the activity of thrombotic factors and inhibit fibrinolysis (6). Data from the literature suggests that hyperuricemia may predispose individuals to a higher incidence of thromboembolic events (7,8).

Albumin is an acute phase reactive protein, and it has several essential functions, such as osmotic pressure maintenance, transport, and regulation of pH. Serum albumin has various protective mechanisms in the vascular bed. It has antioxidant properties, improves endothelial function, and inhibits platelet aggregation via neutralizing reactive oxygen and nitrogen species. Furthermore, albumin binds and activates antithrombin, inhibits factor Xa, and decreases the hepatic synthesis of factors V and VII. Meta-analyses have demonstrated a significant association between hypoalbuminemia and venous thromboembolism (VTE) (9,10).

Increased UA and decreased serum albumin are associated with prothrombotic and inflammatory states and could contribute to VTE pathogenesis. Given the established roles of UA and albumin in thrombosis, we explored whether their ratio could predict 30-day all-cause mortality in acute PE.

Materials and Methods

This retrospective study initially assessed 320 adult patients diagnosed with acute PE and admitted to our center between June 2011 and August 2024. We obtained data on demographics, clinical status, and laboratory

results by examining both written medical files and the hospital database. Acute PE diagnosis was confirmed by CTPA. In patients with hemodynamic instability and no absolute contraindications for thrombolysis, systemic fibrinolysis by infusing 100 mg of recombinant tissue-type plasminogen activator over two hours was administered. In hemodynamically unstable patients for whom thrombolytic therapy was an absolute contraindication, percutaneous catheter-directed interventions were performed as reperfusion therapy. Initial anticoagulation for all patients was provided with either low-molecular-weight heparin or a direct oral anticoagulant, specifically apixaban or rivaroxaban. For a proportion of the patients, long-term oral anticoagulation was achieved with a vitamin K antagonist, following a bridging period with low-molecular-weight heparin.

Comorbidities were defined according to established diagnostic criteria: Cerebrovascular accident (11), coronary artery disease (CAD) (12), hypertension (13), diabetes mellitus (14), chronic kidney disease (CKD) (15), chronic obstructive pulmonary disease (16), and heart failure (HF) (17). The estimated glomerular filtration rate (eGFR) was derived from the chronic kidney disease epidemiology collaboration (CKD-EPI) equation. Patients who died from any cause in the first month of follow-up were labeled as non-survivors, and the remaining as survivors.

After excluding patients with active infection, severe CKD (eGFR <30 mL/min/1.73 m²) (15), inflammatory or hematological disorders, a history of gout or use of urate-lowering therapy, significant hepatic dysfunction (irreversible hepatic dysfunction, manifested by ascites, hepatic encephalopathy, variceal bleeding, or jaundice) (18), or incomplete datasets, the final analysis had 261 patients (Figure 1).

Transthoracic echocardiography (Vivid S70; GE Medical System, Horten, Norway) was carried out to assess the cardiac function, valve morphology, systolic pulmonary artery pressure (sPAP), tricuspid annular plane systolic excursion (TAPSE), and cardiac chamber dimensions. Left ventricular ejection fraction (LV-EF) was measured using Simpson's method, and sPAP was estimated from the Doppler-derived peak tricuspid regurgitation gradient, to which an estimated right atrial pressure was added. Right atrial pressure was inferred from inferior vena cava diameter and its respiratory variation during a forced inspiratory maneuver. The right ventricle-to-left ventricle diameter (RV/LV) ratio was determined using CTPA. Right and left ventricular short-axis dimensions were measured in optimal four-chamber views, just below the respective atrioventricular valve.

We obtained venous blood specimens from each patient at baseline, prior to starting treatment. Measurements were conducted using a BS-2000M chemistry analyzer (Mindray, Shenzhen, China). The uric acid-to-albumin ratio (UAR) was calculated as the ratio of baseline serum UA (mg/dL) to baseline serum albumin (g/dL). We calculated original PESI scores using the established criteria summarized in Table 1 (4). Patients were divided into low-risk, moderate/high-risk, and very high-risk groups based on their PESI scores.

Table 1. Original pulmonary embolism severity index

Variables		1-month mortality risk
Age	Years	≤65 points: Very low (0-1.6%)
Male sex	+10 points	66-85 points: Low (1.7-3.5%)
Cancer	+30 points	86-105 points: Moderate (3.2-7.1%)
Chronic heart failure	+10 points	106-125 points: High (4-11.4%)
Chronic pulmonary disease	+10 points	>125 points: Very high (10-24.5%)
Heart rate ≥110 b.p.m.	+20 points	
SBP <100 mmHg	+30 points	
Respiratory rate >30 per min	+20 points	
Temperature <36 °C	+20 points	
Altered level of consciousness	+60 points	
Arterial oxygen saturation <90%	+20 points	

b.p.m.: Beats per minute, SBP: Systolic blood pressure

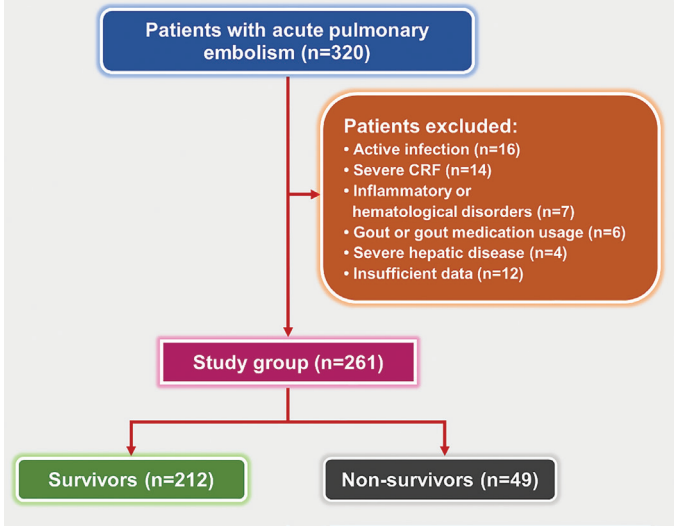


Figure 1. Flow diagram for the study
CRF: Chronic renal failure

The study received approval from the Local Ethics Committee of University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital (decision number: 2025/10/15/099, date: 24.10.2025) and was executed in line with the Declaration of Helsinki. In our study, data from patients between June 2011 and August 2024 were analyzed retrospectively.

Statistical Analysis

Descriptive statistics, group comparisons, and regression analyses were performed using SPSS version 26.0 (IBM Corp., Armonk, NY, USA). Receiver operating characteristic (ROC) curve analysis, including AUC estimation and comparison of correlated ROC curves using the DeLong method, were conducted in R statistical software (R Foundation for Statistical Computing, Vienna, Austria) with the pROC package. Nominal variables are represented as counts and percentages. Normality of continuous data was examined through the Kolmogorov-Smirnov test, and results were reported as the mean ± standard deviation or median with interquartile range. Differences in frequency distributions were tested with the Pearson chi-square test, the chi-square test with Yates' continuity correction, or Fisher's exact test, as indicated. Continuous variables were evaluated with the Mann-Whitney U test or Student's t-test, after assessing their distribution. Logistic regression analysis was employed to evaluate the impact of independent predictors on mortality. The multivariable model was built by incorporating variables that emerged as significant in the preceding univariable analysis. The two-tailed p-value <0.05 was considered statistically significant.

Results

After applying inclusion criteria, 261 patients with acute PE were retained for analysis. The patients' mean age was 63.6±15.9 years, with a female predominance of 60.2%. During the 30-day follow-up period, mortality occurred in 49 patients (18.8%). The average PESI score was 131.2±45.3.

Reperfusion therapy was administered to 77 of the 80 high-risk patients via intravenous rtPA treatment and to the three patients with thrombolytic contraindications via percutaneous catheter-directed intervention.

Survivors and non-survivors exhibited no statistically meaningful differences regarding sex, CAD, diabetes mellitus, cerebrovascular accident, hypertension, chronic obstructive pulmonary disease, HF, deep vein thrombosis, LV-EF, TAPSE, high-sensitivity cardiac troponin I (hs-cTnI), D-dimer, hemoglobin, and platelet count. Non-survivors had significantly higher age (69.5±16.4 vs. 62.3±15.6; p=0.004), PESI score (183.5±53.8 vs. 119.1±33.1; p<0.001), malignancy history (28.6% vs. 8.5%; p<0.001), sPAP (50.6±11.4 vs. 46.3±11.3; p=0.016), RV/LV ratio >1 (77.6% vs. 60.8%; p=0.042), UA (7.8±2.8 vs. 5.8±1.8; p<0.001), and UAR (2.5±1.1 vs. 1.5±0.5; p<0.001). However, eGFR (59.3±23 vs. 80±24.5; p<0.001) and albumin (3.3±0.6 vs. 3.9±0.5; p<0.001) were significantly lower among non-survivors (Table 2).

ROC curve analysis demonstrated that the UAR had a moderate discriminatory ability for predicting 30-day mortality in the overall cohort. The area under the curve (AUC) was 0.788 (95% CI: 0.716-0.860). The optimal cut-off value was 1.78, with a sensitivity of 71.4% and a specificity of 67.9% for predicting 30-day mortality. In the low-risk group (n=10), no 30-day mortality events were observed; therefore, ROC analysis could not be performed for this subgroup. In the moderate/high-risk group (n=171; 25 deaths), UAR demonstrated moderate discriminatory ability for predicting 30-day mortality, with an AUC of 0.753 (95% CI: 0.642-0.863). The optimal cut-off value was 1.7, yielding a sensitivity of 68% and a specificity of 67.1%. In the very high-risk group (n=80; 24 deaths), UAR also showed moderate predictive performance, with an AUC of 0.798 (95% CI: 0.691-0.906). The optimal cut-off value was 1.84, corresponding to a sensitivity of 66.7% and a specificity of 66.1% (Figure 2). Comparison of AUCs between the

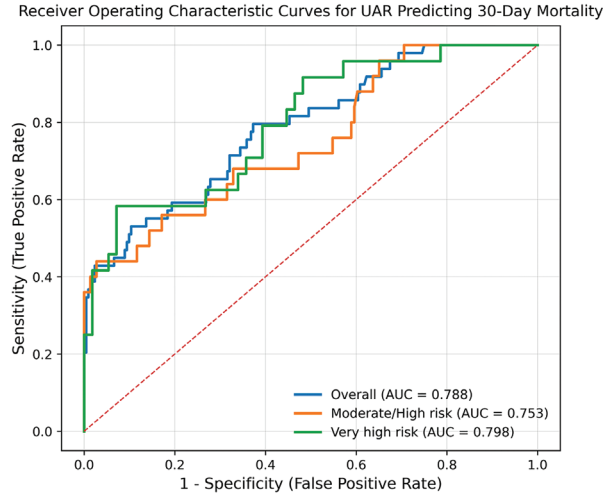


Figure 2. Receiver operating characteristic (ROC) curves of the uric acid-to-albumin ratio (UAR) for predicting 30-day mortality in the overall cohort and risk subgroups. UAR demonstrated moderate discriminative ability in the overall population (AUC =0.788), as well as in the moderate/high-risk (AUC =0.753) and very high-risk groups (AUC =0.798)

AUC: Area under the curve

moderate/high and very high-risk groups using DeLong's method did not show a statistically significant difference (ΔAUC =0.046; p=0.562).

Patients were categorized into low-UAR (<1.78, n=161) and high-UAR (≥1.78, n=100) groups. High-UAR patients exhibited significantly higher age (70.8±13.7 vs. 59.2±15.7; p<0.001), PESI score (147.8±46 vs. 120.8±41.8; p<0.001), hs-cTnI levels [86.6 (38.5-250) vs. 52.5 (18-182); p=0.022], and prevalence of CAD (21% vs. 7.5%; p=0.003), hypertension (59% vs. 40.4%; p=0.005), and HF (11% vs. 1.2%; p=0.001), along with lower LV-EF [55 (55-60) vs. 60 (55-60); p=0.004], TAPSE (14.2±3 vs. 15.3±3.1; p=0.006), and eGFR (61±23.5 vs. 85.5±22.3; p<0.001) (Table 3).

Univariable and multivariable logistic regression models were constructed to identify factors independently associated with 30-day mortality. Age, history of malignancy, right ventricle/left ventricle (RV/LV) ratio >1, sPAP, hs-cTnI, and UAR were each positively associated, whereas eGFR was negatively associated with mortality in the univariable analysis. A high degree of correlation was observed between RV/LV ratio and sPAP; therefore, to prevent collinearity in the regression model, only RV/LV ratio was entered into the multivariable analysis. Afterwards, these factors were evaluated with a multivariable model. The analysis indicated that malignancy history (OR 3.343; 95% CI 1.218-9.170; p=0.019) and high UAR (OR 4.798; 95% CI 2.485-9.262; p<0.001) independently predicted mortality within 30-day (Table 4).

Table 2. Baseline characteristics of patients in terms of 30-day mortality

Variables	Total population (n= 261)	Survivors (n= 212)	Non-survivors (n= 49)	p-value
Comorbidity and clinical properties				
Age, (years)	63.6±15.9	62.3±15.6	69.5±16.4	0.004
Female sex, n (%)	157 (60.2)	127 (59.9)	30 (61.2)	0.994
CAD, n (%)	33 (12.6)	26 (12.3)	7 (14.3)	0.701
DM, n (%)	54 (20.7)	46 (21.7)	8 (16.3)	0.522
Hypertension, n (%)	124 (47.5)	104 (49.1)	20 (40.8)	0.378
CVA, n (%)	14 (5.4)	9 (4.2)	5 (10.2)	0.149
COPD, n (%)	31 (11.9)	26 (12.3)	5 (10.2)	0.875
Heart failure, n (%)	13 (5)	9 (4.2)	4 (8.2)	0.274
Malignancy history, n (%)	32 (12.3)	18 (8.5)	14 (28.6)	**<0.001**
DVT, n (%)	75 (28.7)	60 (28.3)	15 (30.6)	0.747
PESI score	131.2±45.3	119.1±33.1	183.5±53.8	**<0.001**
Echocardiographic parameters				
• LV-EF (%)	58 [55-60]	60 [55-60]	57 [55-60]	0.239
• sPAP, (mmHg)	47.1±11.3	46.3±11.3	50.6±11.4	**0.016**
• TAPSE, (mm)	14.9±3.1	15.1±3.1	14.2±2.9	0.061
• RV/LV ratio >1, n (%)	167 (64)	129 (60.8)	38 (77.6)	**0.042**
Laboratory				
• hs-cTnI, (pg/mL)	64.5 [25-215]	58.7 [24.2-201]	89.9 [28.9-300]	0.146
• D-dimer, (ng/mL)	4.4 [1.8-7.9]	4.1 [1.8-7.8]	5.4 [1.1-8.2]	0.807
• eGFR, (mL/min/1.73 m^2)	76.1±25.7	80±24.5	59.3±23	**<0.001**
• Uric acid, (mg/dL)	6.1±2.2	5.8±1.8	7.8±2.8	**<0.001**
• Albumin, (g/dL)	3.7±0.6	3.9±0.5	3.3±0.6	**<0.001**
• Hemoglobin, (gr/dL)	12.4±1.9	12.5±1.9	12±2.2	0.160
• Platelet, (10^3 µL)	228.5±83.6	224.8±79.9	244.5±97.4	0.137
• UAR	1.7±0.8	1.5±0.5	2.5±1.1	**<0.001**

CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, CVA: Cerebrovascular accident, DM: Diabetes mellitus, DVT: Deep venous thrombosis, eGFR: Estimated glomerular filtration rate, EF: Ejection fraction, hs-cTnI: High-sensitivity cardiac troponin I, LV: Left ventricle, PESI: Pulmonary embolism severity index, RV/LV: Right ventricle-to-left ventricle diameter, sPAP: Systolic pulmonary artery pressure, TAPSE: Tricuspid annular plane systolic excursion, UAR: Uric acid-to-albumin ratio. P-value <0.05 was regarded as statistically significant

Table 3. Baseline characteristics of patients grouped by UAR

Variables	Total population (n=261)	Low-UAR (n=161)	High-UAR (n=100)	p-value
Comorbidity and clinical properties				
Age, (years)	63.6±15.9	59.2±15.7	70.8±13.7	**<0.001**
Female sex, n (%)	157 (60.2)	92 (57.1)	65 (65)	0.258
CAD, n (%)	33 (12.6)	12 (7.5)	21 (21)	**0.003**
DM, n (%)	54 (20.7)	38 (23.6)	16 (16)	0.188
Hypertension, n (%)	124 (47.5)	65 (40.4)	59 (59)	0.005
CVA, n (%)	14 (5.4)	7 (4.3)	7 (7)	0.355
COPD, n (%)	31 (11.9)	20 (12.4)	11 (11)	0.882
Heart failure, n (%)	13 (5)	2 (1.2)	11 (11)	**0.001**
Malignancy history, n (%)	32(12.3)	17 (10.6)	15 (15)	0.288
DVT, n (%)	75 (28.7)	49 (30.4)	26 (26)	0.529
PESI score	131.2±45.3	120.8±41.8	147.8±46	**<0.001**
Echocardiographic parameters				
• LV-EF (%)	58 [55-60]	60 [55-60]	55 [55-60]	**0.004**
• sPAP, (mmHg)	47.9±11.7	46.3±11.7	48.4±10.7	0.156
• TAPSE, (mm)	14.9±3.1	15.3±3.1	14.2±3	**0.006**
• RV/LV ratio >1, n (%)	167 (64)	96 (59.6)	71 (71)	0.084
Laboratory				
• hs-cTnI, (pg/mL)	65.5 [25-215]	52.5 [18-182]	86.6 [38.5-250]	**0.022**
• D-dimer, (ng/mL)	4.4 [1.8.-7.9]	4.4 [1.6-7.8]	4.4 [1.9-7.9]	0.812
• eGFR, (mL/min/1.73 m^2)	76.1±25.7	85.5±22.3	61±23.5	**<0.001**
• Hemoglobin, (gr/dL)	12.4±1.9	12.6±1.9	12.1±1.9	0.051
• Platelet, (10^3 µL)	228.5±83.6	225.9±81.5	232.7±87.1	0.519

CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, CVA: Cerebrovascular accident, DM: Diabetes mellitus, DVT: Deep venous thrombosis, eGFR: Estimated glomerular filtration rate, EF: Ejection fraction, hs-cTnI: High-sensitivity cardiac troponin I, LV: Left ventricle, PESI: Pulmonary embolism severity index, RV/LV: Right ventricle-to-left ventricle diameter, sPAP: Systolic pulmonary artery pressure, TAPSE: Tricuspid annular plane systolic excursion, UAR: Uric acid to albumin ratio. P-value <0.05 was regarded as statistically significant

Discussion

This retrospective study evaluated the prognostic value of the UAR in patients with acute PE. ROC curve analysis showed that UAR has a significant predictive performance for 30-day mortality in the acute PE population. Elevated UAR and the history of malignant disease were identified as independent predictors of mortality within 30 days. This is the first investigation, to the best of our knowledge, to assess the prognostic value of the UAR in this patient population.

PE remains one of the contributing factors to high mortality rates from cardiovascular diseases. Although clinical prediction models such as PESI and simplified PESI are widely used, they rely primarily on demographic and clinical variables. Also, in patients with acute PE, various diagnostic tools—including cardiac troponins, CTPA, and transthoracic echocardiography—have been employed to assess right ventricular dysfunction and predict short-term outcomes. Additionally, emerging laboratory biomarkers such as H-FABP, NT-proBNP, and copeptin have been shown to have further prognostic information in observational cohort studies. While their role in guiding therapeutic strategies has not been validated through randomized controlled trials yet (4). These scores do not fully reflect some of the mechanisms, such as oxidative stress, inflammation, and endothelial injury, that may affect the prognosis in these patients.

UA contributes to vascular pathology through multiple mechanisms. It is not only an end-product of purine metabolism but also a pro-oxidant molecule that can stimulate inflammation and exacerbate endothelial dysfunction. Evidence suggests that high UA levels play a role in the stimulation of the renin-angiotensin-aldosterone system, increased oxidative stress, smooth muscle cell proliferation, and prothrombotic states (6,7,19-22). VTE is linked to the activation of both inflammatory

Table 4. Predictors of 30-day mortality in acute pulmonary embolism patients

Variables	Univariable OR (95% CI)	p-value	Multivariable OR (95% CI)	p-value
Age	1.03 (1.01-1.06)	0.005	0.997 (0.968-1.026)	0.822
Female sex	0.865 (0.501-1.789)	0.865		
CAD	1.192 (0.485-2.93)	0.701		
Diabetes mellitus	0.704 (0.309-1.607)	0.405		
Hypertension	0.761 (0.381-1.345)	0.299		
CVA	2.563 (0.819-8.02)	0.106		
COPD	0.813 (0.296-2.236)	0.688		
Heart failure	2.005 (0.591-6.8)	0.418		
Malignancy history	4.311 (1.965-9.459)	**<0.001**	3.343 (1.218-9.170)	**0.019**
DVT	1.118 (0.568-2.2)	0.747		
LV-EF	0.953 (0.895-1.014)	0.125		
sPAP	1.033 (1.006-1.062)	**0.018**	1.012 (0.978-1.047)	0.499
TAPSE	0.905 (0.814-1.005)	0.063		
RV/LV ratio >1	2.223 (1.076-4.592)	**0.031**		
eGFR	0.966 (0.952-0.984)	**<0.001**	0.989 (0.971-1.007)	0.236
Baseline hemoglobin	0.892 (0.76-1.046)	0.160		
Baseline platelet	1.003 (0.999-1.006)	0.139		
hs-cTnI	1.001 (1.00003-1.001)	**0.038**	1.000 (1.000-1.001)	0.319
D-dimer	1.017 (0.934-1.107)	0.700		
UAR	5.557 (3.21-9.688)	**<0.001**	4.798 (2.485-9.262)	**<0.001**

CAD: Coronary artery disease, COPD: Chronic obstructive pulmonary disease, CVA: Cerebrovascular accident, DVT: Deep venous thrombosis, EF: Ejection fraction, eGFR: Estimated glomerular filtration rate, hs-cTnI: High-sensitivity cardiac troponin I, LV: Left ventricle, RV/LVR: Right ventricle-to-left ventricle diameter, sPAP: systolic pulmonary artery pressure, TAPSE: Tricuspid annular plane systolic excursion, UAR: Uric acid to albumin ratio. P-value <0.05 was regarded as statistically significant

and coagulation pathways (23), and imbalances in these systems may increase the risk of acute PE. The Atherosclerosis Risk in Communities study, involving 14,126 participants, demonstrated a significant association between serum UA and VTE incidence (7). Moreover, 280 patients with previous VTE were evaluated in another study, and a significant association has also been identified between serum UA levels and the risk of recurrent VTE (24). Furthermore, among 337 patients with verified PE, those in the intermediate- and high-risk categories exhibited higher UA levels, which independently predicted mortality within 30 days (25).

Albumin is an indicator of both nutritional status and systemic inflammation. Owing to its negative acute-phase characteristics, systemic inflammation and oxidative stress cause a decrease in its concentration. Albumin exhibits anti-inflammatory, antioxidant, and anticoagulant effects by neutralizing reactive oxygen species, enhancing endothelial stability, and reducing platelet aggregation (26-31). Hypoalbuminemia has been associated with hypercoagulability, increased blood viscosity, and impaired vascular integrity, which all predispose to thrombotic complications (9). Evidence from meta-analyses reveals that hypoalbuminemia is a significant risk factor for VTE (9,10).

The UAR combines two biomarkers with opposite biological effects: UA has oxidative, inflammatory, and thrombogenic properties, whereas albumin has anti-inflammatory, antioxidant, and anti-thrombogenic capacity. Hence, UAR may more accurately reflect the overall pathophysiological processes than either marker alone. Recent studies have demonstrated that UAR independently predicts mortality in several cardiovascular diseases, such as acute myocardial infarction (32), unstable angina (33), and acute aortic dissection (34). We found UAR as an independent predictor of 30-day mortality. Our findings support the fact that systemic antithrombotic, antioxidant, and anti-inflammatory dysregulation plays a key role in physiopathology and the prognosis of PE.

UAR can be calculated easily using routine laboratory parameters that are inexpensive and widely available, without requiring specialized assays or additional costs. Incorporating UAR into existing risk assessment models

could enhance their early mortality predictive accuracy, particularly in patients with intermediate to high risk.

Study Limitations

This study has various limitations. First, it is a retrospective and single-center investigation that limits the ability to infer definitive causal conclusions and carries the potential for selection and information biases. Although consecutive patients were evaluated, the generalizability of the results may be affected due to the restriction of the final analysis to patients with complete data. Furthermore, statistical power may be reduced despite multivariable adjustment because of modest sample size and limited number of events that increase the risk of residual confounding. Second, the findings were confined to short-term mortality in acute PE patients, and long-term outcomes were not evaluated. Third, although patients with active infection, severe CKD, significant hepatic dysfunction, inflammatory or hematological disorders, gout, or urate-lowering therapy were excluded, several other conditions that may influence serum UA and albumin levels were not systematically assessed. Serum albumin levels may be affected by nutritional status, nephrotic proteinuria, protein-losing enteropathy, volume overload, or acute stress states. Besides, serum UA levels can be influenced by dietary factors, obesity, thyroid dysfunction, diseases that cause increased cell turnover, subclinical renal diseases, dehydration, genetic factors, and medications that are not included in our analysis, such as levodopa, theophylline, and anabolic steroids. The lack of detailed investigation of these potential confounding factors may have affected the observed results. Fourth, only baseline values of serum UA and albumin were considered, but serial measurements were not performed. Therefore, the prognostic value of the UAR dynamic change was not established. Despite these limitations, our findings highlight UAR as a readily accessible biomarker predicting short-term mortality risk in patients with acute PE.

Conclusion

In this study, the UAR was found to be an independent prognostic factor for 30-day mortality in acute PE. This simple, inexpensive biomarker, which is derived from routine tests, may enhance early risk stratification and guide more individualized management strategies when integrated into existing prognostic models. However, large-scale, multicenter prospective studies are needed to confirm these findings and clarify the clinical utility of UAR in the prognostic assessment of acute PE patients.

Ethics

Ethics Committee Approval: The study received approval from the Local Ethics Committee of University of Health Sciences Turkey, İstanbul Bağcılar Training and Research Hospital (decision number: 2025/10/15/099, date: 24.10.2025).

Informed Consent: In our study, data from patients between June 2011 and August 2024 were analyzed retrospectively.

Footnotes

Authorship Contributions

Surgical and Medical Practices: O.İ., E.D., M.F.D., C.T., S.B., S.Ö., Concept: O.İ., S.Ö., Design: O.İ., S.Ö., Data Collection or Processing: O.İ., E.D., M.F.D., C.T., S.B., S.Ö., Analysis or Interpretation: O.İ., Literature Search: O.İ., M.F.D., C.T., S.B., Writing: O.İ., E.D., S.Ö.

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

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

References

1. Raskob GE, Angchaisuksiri P, Blanco AN, Buller H, Gallus A, Hunt BJ, et al. Thrombosis: a major contributor to global disease burden. Arterioscler Thromb Vasc Biol. 2014;34(11):2363-2371.
2. Iskandar JP, Hariri E, Kanaan C, Kassis N, Kamran H, Sese D, et al. The safety and efficacy of systemic versus catheter-based therapies: application of a prognostic model by a pulmonary embolism response team. J Thromb Thrombolysis. 2022;53(3):616-625.
3. Solomon CG. Pulmonory embolism. N Engl J Med. 2022;387:45-57.
4. Konstantinides SV, Meyer G, Becattini C, Bueno H, Geersing GJ, Harjola VP, et al. 2019 ESC guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS). Eur Heart J. 2020;41(4):543-603.
5. Murguia AR, Segovia F, Ayvali F, Brockman M, Prakash S, Singh V, et al. Evaluation of four validated risk scores to predict outcomes in hispanic patients with acute pulmonary embolism. Angiology. 2025;76(7):629-636.
6. Țăpoi L, Șalaru DL, Sascău R, Stătescu C. Uric acid-an emergent risk marker for thrombosis? J Clin Med. 2021;10(10):2062.
7. Kubota Y, McAdams-DeMarco M, Folsom AR. Serum uric acid, gout, and venous thromboembolism: the atherosclerosis risk in communities study. Thromb Res. 2016;144:144-148.
8. Weng H, Li H, Zhang Z, Zhang Y, Xi L, Zhang D, et al. Association between uric acid and risk of venous thromboembolism in East Asian populations: a cohort and Mendelian randomization study. Lancet Reg Health West Pac. 2023;39:100848.
9. Valeriani E, Pannunzio A, Palumbo IM, Bartimoccia S, Cammisotto V, Castellani V, et al. Risk of venous thromboembolism and arterial

events in patients with hypoalbuminemia: a comprehensive meta-analysis of more than 2 million patients. J Thromb Haemost. 2024;22(10):2823-2833.

10. Seidu S, Kunutsor SK, Khunti K. Serum albumin, cardiometabolic and other adverse outcomes: systematic review and meta-analyses of 48 published observational cohort studies involving 1,492,237 participants. Scand Cardiovasc J. 2020;54(5):280-293.
11. Mansfield A, Inness EL, Mcilroy WE. Stroke. Handb Clin Neurol. 2018;159:205-228.
12. Vrints C, Andreotti F, Koskinas KC, Rossello X, Adamo M, Ainslie J, et al. 2024 ESC guidelines for the management of chronic coronary syndromes. Eur Heart J. 2024;45(36):3415-3537.
13. McEvoy JW, McCarthy CP, Bruno RM, Brouwers S, Canavan MD, Ceconi C, et al. 2024 ESC guidelines for the management of elevated blood pressure and hypertension. Eur Heart J. 2024;45(38):3912-4018.
14. Petersmann A, Müller-Wieland D, Müller UA, Landgraf R, Nauck M, Freckmann G, et al. Definition, classification and diagnosis of diabetes mellitus. Exp Clin Endocrinol Diabetes. 2019;127(S 01):S1-S7.
15. Kidney Disease: Improving Global Outcomes (KDIGO) CKD Work Group. KDIGO 2024 clinical practice guideline for the evaluation and management of chronic kidney disease. Kidney Int. 2024;105(4S):S117-S314.
16. Halpin DMG, Criner GJ, Papi A, Singh D, Anzueto A, Martinez FJ, et al. Global initiative for the diagnosis, management, and prevention of chronic obstructive lung disease. The 2020 GOLD science committee report on COVID-19 and chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2021;203(1):24-36.
17. Heidenreich PA, Bozkurt B, Aguilar D, Allen LA, Byun JJ, Colvin MM, et al. 2022 AHA/ACC/HFSA guideline for the management of heart failure: a report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. J Am Coll Cardiol. 2022;79(17):e263-e421.
18. European Association for the Study of the Liver. EASL clinical practice guidelines on acute-on-chronic liver failure. J Hepatol. 2023;79(2):461-491.
19. Mannarino MR, Pirro M, Gigante B, Savonen K, Kurl S, Giral P, et al. Association between uric acid, carotid intima-media thickness, and cardiovascular events: prospective results from the IMPROVE study. J Am Heart Assoc. 2021;10(11):e020419.
20. Kanbay M, Jensen T, Solak Y, Le M, Roncal-Jimenez C, Rivard C, et al. Uric acid in metabolic syndrome: From an innocent bystander to a central player. Eur J Intern Med. 2016;29:3-8.
21. Jayachandran M, Qu S. Harnessing hyperuricemia to atherosclerosis and understanding its mechanistic dependence. Med Res Rev. 2021;41(1):616-629.
22. Zapolski T, Waciński P, Kondracki B, Rychta E, Buraczyńska MJ, Wysokiński A. Uric acid as a link between renal dysfunction and both pro-inflammatory and prothrombotic state in patients with metabolic syndrome and coronary artery disease. Kardiol Pol. 2011;69(4):319-326.
23. Colling ME, Tourdot BE, Kanthi Y. Inflammation, infection and venous thromboembolism. Circ Res. 2021;128(12):2017-2036.
24. De Lucchi L, Nardin C, Sponchiado A, Raggi D, Faggin E, Martini E, et al. Serum uric acid levels and the risk of recurrent venous thromboembolism. J Thromb Haemost. 2021;19(1):194-201.
25. Ozsu S, Çoşar AM, Aksoy HB, Bülbül Y, Oztuna F, Karahan SC, et al. Prognostic value of uric acid for pulmonary thromboembolism. Respir Care. 2017;62(8):1091-1096.
26. Fanali G, di Masi A, Trezza V, Marino M, Fasano M, Ascenzi P. Human serum albumin: from bench to bedside. Mol Aspects Med. 2012;33(3):209-290.
27. Belayev L, Pinard E, Nallet H, Seylaz J, Liu Y, Riyamongkol P, et al. Albumin therapy of transient focal cerebral ischemia: in vivo analysis of dynamic microvascular responses. Stroke. 2002;33(4):1077-1084.
28. Manolis AA, Manolis TA, Melita H, Mikhailidis DP, Manolis AS. Low serum albumin: a neglected predictor in patients with cardiovascular disease. Eur J Intern Med. 2022;102:24-39.
29. Basili S, Carnevale R, Nocella C, Bartimoccia S, Raparelli V, Talerico G, et al. Serum albumin is inversely associated with portal vein thrombosis in cirrhosis. Hepatol Commun. 2019;3(4):504-512.
30. Belinskaia DA, Voronina PA, Shmurak VI, Vovk MA, Batalova AA, Jenkins RO, et al. The universal soldier: enzymatic and non-enzymatic antioxidant functions of serum albumin. Antioxidants (Basel). 2020;9(10):966.
31. Belinskaia DA, Voronina PA, Goncharov NV. Integrative role of albumin: evolutionary, biochemical and pathophysiological aspects. J Evol Biochem Physiol. 2021;57(6):1419-1448.
32. Kalkan S, Cagan Efe S, Karagöz A, Zeren G, Yılmaz MF, Şimşek B, et al. A new predictor of mortality in ST-elevation myocardial infarction: the uric acid albumin ratio. Angiology. 2022;73(5):461-469.
33. Li S, Chen H, Zhou L, Cui H, Liang S, Li H. The uric acid to albumin ratio: a novel predictor of long-term cardiac mortality in patients with unstable angina pectoris after percutaneous coronary intervention. Scand J Clin Lab Invest. 2022;82(4):304-310.
34. Wang X, Deng C, Guo F, Zhong L, Gao H. The preoperative uric acid-to-albumin ratio as a new indicator to predict long-term prognosis after surgery for patients with acute type A aortic dissection. Heart Surg Forum. 2023;26(1):E001-E008.